

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:42:12 ; Search time 55.68 Seconds  
(without alignments)  
783.980 Million cell updates/sec

Title: US-09-147-052-2\_COPY\_64\_456

Perfect score: 1998

Sequence: 1 CWSITKDKANPNNGQTLEA.....SSNENNADKIPCYRRPGTFL 393

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1998	100.0	456	18 AAW36050	Hybrid Marek's dis
2	1994	99.8	1086	18 AAW36051	Hybrid Marek's dis
3	1910	95.6	615	15 AAR63230	Mycoplasma gallise
4	1856	92.9	610	15 AAR63229	Mycoplasma gallise
5	1612	80.7	368	14 AAR44493	Mycoplasma gallise
6	1612	80.7	368	15 AAR63227	Mycoplasma gallise
7	1604	80.3	368	16 AAR76955	Mycoplasma gallise
8	1142	57.2	235	10 AAR93646	Amino acid sequenc
9	1142	57.2	235	11 AAR05081	MG-1 antigen. AA
10	1142	57.2	261	11 AAR05082	TMG-1 antigen. A
11	1142	57.2	261	16 AAR79911	M.gallisepticum 26

12	1117	55.9	261	10 AAR93959	Amino acid (AA) se
13	806	40.3	661	15 AAR63226	Mycoplasma gallise
14	806	40.3	661	16 AAR79910	M.gallisepticum 66
15	713	35.7	647	16 AAW11978	Mycoplasma gall
16	700.5	35.1	648	15 AAR56973	PMGA 1.2 protein o
17	386.5	19.3	183	10 AAR93649	Amino acid sequenc
18	386.5	19.3	183	11 AAR06439	MG-4 antigen. AA
19	284.5	14.2	219	10 AAR93648	Amino acid sequenc
20	284.5	14.2	219	11 AAR06438	MG-3 antigen. AA
21	176.5	8.8	6281	22 AAW37403	Staphylococcus aur
22	176	8.8	1095	22 AAG83030	S. epidermidis ope
23	171	8.6	2086	22 AAW34143	Staphylococcus aur
24	171	8.6	5795	22 AAW37017	Staphylococcus aur
25	162	8.1	1029	22 AAW34389	Staphylococcus aur
26	162	8.1	1048	22 AAW37490	Staphylococcus aur
27	159	8.0	933	21 AAY58435	Staphylococcus aur
28	159	8.0	933	22 AAW69508	Staphylococcus aur
29	159	8.0	936	18 AAW89801	Staphylococcus aur
30	158	7.9	2434	22 AAW34339	Staphylococcus aur
31	156	7.8	807	21 AAB18311	Plasmodium falcipa
32	154.5	7.7	682	17 AAR95273	Nisin nisp gene pr
33	153.5	7.7	2478	22 AAW34320	Staphylococcus aur
34	153.5	7.7	2478	22 AAW37374	Staphylococcus aur
35	151.5	7.6	1237	18 AAW55640	H. pylori ORF 04ep
36	151.5	7.6	1237	20 AAW17187	H. pylori outer me
37	151	7.6	5024	22 AAG82935	S. epidermidis ope
38	145.5	7.3	2437	22 AAW34338	Staphylococcus aur
39	145	7.3	1215	22 AAW34412	Staphylococcus aur
40	145	7.3	1269	22 AAW37520	Staphylococcus aur
41	143.5	7.2	837	22 AAW34387	Staphylococcus aur
42	143.5	7.2	875	22 AAW37487	Staphylococcus aur
43	143	7.2	3158	22 AAW37018	Staphylococcus aur
44	142	7.1	135	15 AAR56974	Partial PMGA 1.3 p
45	142	7.1	135	16 AAW11979	Mycobacterium gall

#### ALIGNMENTS

RESULT 1

AAW36050  
AAW36050 standard; Protein; 456 AA.

XX AC AAW36050;

XX XX  
DT 15-JUL-1998 (first entry)

XX XX Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX DE Chimeric; Marek's disease virus; outer membrane protein; fusion protein;  
antigen; vaccine; poultry.

XX OS Chimeric - Marek's disease gammaherpesvirus.  
OS OS Chimeric - Mycoplasma gallisepticum.

XX XX Key Location/Qualifiers  
XX XX Region 1..64

XX XX Region 65..456  
FT FT /note= "derived from Marek's disease virus gB protein"  
FT FT /note= "derived from M. gallisepticum antigenic protein"

XX XX WO9736924-A1.  
XX XX 09-OCT-1997.

XX XX 28-MAR-1997; 97WO-JF01084.

XX XX 29-MAR-1996; 96JP-0103548.

XX XX (JAPG ) NIPPON ZEON KK.

XX XX Saito S, Tsuzaki Y, Yanagida N;

DR WPI; 1997-503046/46.  
 XX N-PSDB; AAT96595.  
 XX Fusion protein comprising herpes virus outer membrane protein and  
 PT antigenic polypeptide - for prevention of infection by Mycoplasma  
 PT gallisepticum, especially in poultry  
 XX  
 PS Disclosure; Page 16-19; 51pp; Japanese.  
 XX  
 XX This sequence represents the chimeric protein 40 K-S which comprises a  
 CC fragment of the Marek's disease virus outer membrane protein gB fused  
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
 CC protein can be used in recombinant live vaccines for prevention of  
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
 CC protein shows antigenicity in poultry.  
 XX  
 XX Sequence 456 AA;  
 SQ  
 Query Match 100.0%; Score 1998; DB 18; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-126;  
 Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CMSITKDNPNNGQQLAARMEITDLINAKAMTILASLDYAKIEASISSAYSEAEVYN 60  
 DB 64 cmsitkdnppnggtgleaarmeltldlinakamtllasiqdyakieasissayseaeavn 123  
 QY 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120  
 DB 124 nlnnatleqlkmaaktulesaingantdkttfdnehpnulveaykalkttleqatnleqls 183  
 QY 121 STAYNIRNVLVLYNKASLIKTLDPLNGGTLDSNEITTANKNINNTLSTINEQKTN 180  
 DB 184 staynqirnlvlynkassliktldpninggtllgsneittanknoinntlstineqktn 243  
 QY 181 ADALSNFVKKVIQNNQSFVGTFTNANVQPSNYSFVPSADVTPVNYKYARTVWNGDE 240  
 DB 244 adalsnfikkvigqnnegsfvgtftnanvqpsnyfsadvtpvnykyartvwnge 303  
 QY 241 PSSRILANTNSITDVSWIYSLAGTNTKYQFSFNGYGPSTGYLYFPYKLVKAADANNVGLQ 300  
 DB 304 pssrillantsitdsvwiyslagntkyqfsfngygpstgylyfpykylvkaadannvqlq 363  
 QY 301 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 360  
 DB 364 yklngnvqqvefatstansnttanptpavdeikvakivlslgrfgntielsvptgegn 423  
 QY 361 MNKVAPMIGNIYLSNENNAADKIPGYRRPCTFL 393  
 DB 424 mnkvapmigniylyssnennadkpggyrrpctfl 456  
 RESULT 2  
 AAW36051  
 ID AAW36051 standard; Protein; 1086 AA.  
 XX  
 XX AAW36051;  
 XX  
 XX 15-JUL-1998 (first entry)  
 XX  
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.  
 XX  
 XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;  
 KW antigen; vaccine; poultry.  
 XX  
 OS Chimeric - Marek's disease gammaherpesvirus.  
 OS Chimeric - Mycoplasma gallisepticum.  
 XX  
 XX Key Location/Qualifiers  
 FT 1..672  
 FT Region /note= "derived from Marek's disease virus gB protein"  
 FT Region 693..1086  
 FT /note= "derived from M. gallisepticum antigen"

XX WO9736924-A1.  
 XX  
 XX 09-OCT-1997.  
 XX 28-MAR-1997; 97WO-JP01084.  
 XX  
 XX 29-MAR-1996; 96JP-0103548.  
 XX  
 XX (JAPG ) NIPPON ZEON KK.  
 XX Saito S, Tsuzaki Y, Yanagida N;  
 PI WPI; 1997-503046/46.  
 XX  
 DR N-PSDB; AAT96596.  
 XX  
 XX Fusion protein comprising herpes virus outer membrane protein and  
 PT antigenic polypeptide - for prevention of infection by Mycoplasma  
 PT gallisepticum, especially in poultry  
 XX  
 PS Disclosure; Page 22-30; 51pp; Japanese.  
 XX  
 XX This sequence represents the chimeric protein 40 K-C which comprises a  
 CC fragment of the Marek's disease virus outer membrane protein gB fused  
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
 CC protein can be used in recombinant live vaccines for prevention of  
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
 CC protein shows antigenicity in poultry.  
 XX  
 XX Sequence 1086 AA;  
 SQ  
 Query Match 99.8%; Score 1994; DB 18; Length 1086;  
 Best Local Similarity 99.7%; Pred. No. 1.2e-125;  
 Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CMSITKDNPNNGQQLAARMEITDLINAKAMTILASLDYAKIEASISSAYSEAEVYN 60  
 DB 694 cmsitkdnppnggtgleaarmeltldlinakamtllasiqdyakieasissayseaeavn 753  
 QY 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120  
 DB 754 nlnnatleqlkmaaktulesaingantdkttfdnehpnulveaykalkttleqatnleqls 813  
 QY 121 STAYNIRNVLVLYNKASLIKTLDPLNGGTLDSNEITTANKNINNTLSTINEQKTN 180  
 DB 814 staynqirnlvlynkassliktldpninggtllgsneittanknoinntlstineqktn 873  
 QY 181 ADALSNFVKKVIQNNQSFVGTFTNANVQPSNYSFVPSADVTPVNYKYARTVWNGDE 240  
 DB 874 adalsnfikkvigqnnegsfvgtftnanvqpsnyfsadvtpvnykyartvwnge 933  
 QY 241 PSSRILANTNSITDVSWIYSLAGTNTKYQFSFNGYGPSTGYLYFPYKLVKAADANNVGLQ 300  
 DB 934 pssrillantsitdsvwiyslagntkyqfsfngygpstgylyfpykylvkaadannvqlq 993  
 QY 301 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 360  
 DB 994 yklngnvqqvefatstansnttanptpavdeikvakivlslgrfgntielsvptgegn 1053  
 QY 361 MNKVAPMIGNIYLSNENNAADKIPGYRRPCTFL 393  
 DB 1054 mnkvapmigniylyssnennadkpggyrrpctfl 1086  
 RESULT 3  
 AAR63230  
 ID AAR63230 standard; Protein; 615 AA.  
 XX  
 XX AAR63230;  
 XX  
 XX 23-JUN-1995 (first entry)  
 XX

```

DE Mycoplasma gallisepticum antigen (UM-67).
KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX Mycoplasma gallisepticum.
OS
XX
XX Key Location/Qualifiers
XX Protein 1..615
XX /note= "Trp residues correspond to TGA codons"
XX
XX WO9423019-A.
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77857.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 87-91; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-67 containing an open reading frame was sequenced (AAQ77857). The
XX ORF encodes an antigenic polypeptide (AAR63230). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX
XX Sequence 615 AA;

Query Match 95.6%; Score 1910; DB 15; Length 615;
Best Local Similarity 98.2%; Pred. No. 2.5e-120;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CMSITKDDANPNNGQTGLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
Db 27 cmsitkddanpnnggtqlqaarmeltdlinakartlasldqyakieaslsayseaeetvn 86

Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 nlnlatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146

Qy 121 STAYNQIRNNLVLYNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180
Db 147 staynqirnnlvlynasslittktdplnggmlldsneltctvnrlnintstineqktn 206

Qy 181 ADALSNSFIKKVQIONNEQSFYGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGDE 240
Db 207 adalnsfikkvqionneqsfygtftnannvqpsnysfvafsadvtvpnykyarrtwngde 266

Qy 241 PSSRLANTNSITDVSWYISLAGNTKYQFSFNGYSTGYLYPPYKLVKAADANNVGLQ 300
Db 267 pssrlantnsitdvswwyislagntkyqfsfngystgylyfpykvlvkaadannvqlq 326

Qy 301 YKLNNGNQVQEEFATSTSANNTTANPTPAVDKIKVAKIVLSGLRFGQNTIELSVPTGEGN 360
Db 327 yklnngnvqvefatstsannttanptpavdeikvakivlsgrlrfgqntielsvptgegn 386

Qy 361 MNKVAPMIGNIYLSNENNADKI 383
Db 383 mknvapmigniylyssennadki

```

```

Db 387 mnkvapmigniylyssennadki 409

RESULT 4
AAR63229
ID AAR63229 standard; Protein: 610 AA.
XX
XX AAR63229;
XX
XX 23-JUN-1995 (first entry)
XX
XX Mycoplasma gallisepticum antigen (UM-66).
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX Mycoplasma gallisepticum.
XX
XX Key Location/Qualifiers
XX Protein 1..610
XX /note= "Trp residues correspond to TGA codons"
XX
XX WO9423019-A.
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77856.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 78-81; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-66 containing an open reading frame was sequenced (AAQ77856). The
XX ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX
XX Sequence 610 AA;

Query Match 92.9%; Score 1856; DB 15; Length 610;
Best Local Similarity 95.5%; Pred. No. 1.1e-116;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CMSITKDDANPNNGQTGLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
Db 27 cmsitkddanpnnggtqlqaarmeltdlinakartlasldqyakieaslsayseaeetvn 86

Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 nlnlatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnleqls 146

Qy 121 STAYNQIRNNLVLYNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180
Db 147 staynqirnnlvlynkasslittktdplnggmlldsneltctvnrlnintstineqktn 206

Qy 181 ADALSNSFIKKVQIONNEQSFYGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGDE 240
Db 207 adalnsfikkvqionneqsfygtftnannvqpsnysfvafsadvtvpnykyarrtwngde 266

```

```

QY 241 PSSRILANTNSITDYSWIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQ 300
  |||
Db 267 psrillantnsitdswiyslsgtntkyqfsfnygstgylyfykvlvkaadannvgliq 326
  |||
QY 301 YKLNGNVOQVEFATSTSANNTTANPTPAVDIKVAKIVLSGLRFGQNTIELSVPTGEGN 360
  |||
Db 327 yklngnvqvfeafatstsannttanptpavdeikvakivlslgrfgqntielsvptgern 386
  |||
QY 361 MNKVAPMIGNIYLSNENNAADK 382
  |||
Db 387 mnkvapmignmytssnaeank 408
  |||

RESULT 5
AAR44493
ID AAR44493 standard; Protein; 368 AA.
XX
AC AAR44493;
XX
DT 16-JUN-1994 (first entry)
XX
DE Mycoplasma gallisepticum 40kD antigen.
XX
KW Vaccine; mycoplasma infection; poultry; fowl.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Misc-difference 262 /note= "corresponds to NNN codon in AAQ53419"
FT FT Misc-difference 283 /note= "corresponds to NNN codon in AAQ53419"
XX
XX WO9324646-A.
XX
XX 09-DEC-1993.
XX
XX 28-MAY-1993; 93WO-JP00715.
XX
XX 29-MAY-1992; 92JP-0138819.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;
XX
XX WPI; 1993-405837/50.
XX N-PSDB; AAQ53419.
XX
XX Mycoplasma gallisepticum antigen and DNA coding for it - useful
XX for vaccination of fowl against mycoplasma infections
XX
XX Claim 2; Page 23-26; 37pp; Japanese.
XX
XX The sequence coding for the 40kDa antigen was obtained by PCR
XX amplification of M.gallisepticum genomic DNA. The antigen reacts
XX with Mycoplasma-immune or Mycoplasma-infected serum and can be used
XX as a vaccine to protect fowl from M.gallisepticum infection.
XX
XX Sequence 368 AA;

Query Match 80.7%; Score 1612; DB 14; Length 368;
Best Local Similarity 95.0%; Pred. No. 1.4e-100;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 CMTSKKDANPNNGQTQLEARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEVTN 60
  |||
Db 27 cmtskkdanpnngqtqiqaarmeltdlinakartlasldqyakieaslsayseaeatvn 86
  |||
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
  |||

```

```

Db 87 nlnnatleqlkmaaktnlesainqantdkttfdnebpnlveaykalkttleqratnlegla 146
QY 121 STAYNQIIRNNLDVLYNKASSLITKTLDPNLGTLTLDNSIEITANKNINNTLSTINEQKTN 180
  |||
Db 147 staynqiirnnldvlynasslittktdplnggmlldnsneittvnrninntlstineqktn 206
  |||
QY 181 ADALNSFTKKVIONEQSFVGTETNANVOPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 240
  |||
Db 207 adalnsftkkvioneqsfvgtetnnavqpsnysfvaafsadvtvpnykyarrtvxngde 266
  |||
QY 241 PSSRILANTNSITDYSWIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQ 300
  |||
Db 267 psrillantnsitdswiyslsgtntkyqfsfnygstgylyfykvlvkaadannvgliq 326
  |||
QY 301 YKLNGNVOQVEFATSTSANNTTANPTPAVDIKVAK 337
  |||
Db 327 yklngnvqvfeafatstsannttanptqqlmrklk 363
  |||

RESULT 6
AAR63227
ID AAR63227 standard; Protein; 368 AA.
XX
AC AAR63227;
XX
DT 23-JUN-1995 (first entry)
XX
DE Mycoplasma gallisepticum 40kD antigen.
XX
KW recombinant avipox virus; live vaccine; mycoplasma 40kD antigen;
XX TTM-1.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Misc-difference 262 /note= "corresponds to a NNN codon"
FT FT Misc-difference 283 /note= "corresponds to a NNN codon"
XX
XX WO9423019-A.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77854.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 71-74; 123pp; Japanese.
XX
XX The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding
XX for the 40kD antigen of Mycoplasma gallisepticum under the control
XX of a synthetic promoter. A 1300 bp restriction fragment containing
XX the promoter-ORF sequence was excised and was used in the
XX construction of plasmid pNZ7929-R2. This in turn was involved in the
XX construction of a recombinant avipox virus vector comprising the
XX TTM-1 gene, DNA encoding the signal membrane anchor peptide from
XX Newcastle Disease Virus haemagglutinin neuraminidase and FVP
XX sequences. The recombinant avipox virus is useful as a live vaccine
XX

```

CC to protect against infection by Mycoplasma gallisepticum.

XX  
SQ Sequence 368 AA;

Query Match 80.7%; Score 1612; DB 15; Length 368;  
Best Local Similarity 95.0%; Pred. No. 1.4e-100;  
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 CMSITKDKANPNNGQTOLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60  
|||||  
Db 27 cmsitkdkanpnngqtlqaarmeltdlinakartlasldyakiessaysaeactn 86  
|||||  
Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 120  
|||||  
Db 87 nlnnatleqlkmaaktlnlesainqantdkttfdnehpnlveaykalkttlegratnleqla 146  
|||||  
Qy 121 STAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTANKNNINNTLSTINEOKTN 180  
|||||  
Db 147 staynqirnnlvlynnasslittktdplnggmildsneittvnrnintlstineqktn 206  
|||||  
Qy 181 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 240  
|||||  
Db 207 adalsnsfikkvionneqsfvgftfnanvqpsnysfvafsdvtpvnykyarrtvxngde 266  
|||||  
Qy 241 PSSRLANTNSITDVSWIYSIAGTNTKYQFSNYGPGSTGYLYFPYKLVKAADANNVGLQ 300  
|||||  
Db 267 pssrlantnsitdvsxiyslagtnkyqfsnygpgstgylyfpyklvkaadannvqlq 326  
|||||  
Qy 301 YKLNGNVQVQVEFATSTANNTTANPTPAVDKIKVAK 337  
|||||  
Db 327 yklngnvqvqvefatstannnttanptqqlmrklk 363  
|||||

## RESULT 7

AAR76955  
ID AAR76955 standard; Protein; 368 AA.

XX  
AC AAR76955;

XX  
DT 12-MAR-1996 (first entry)

XX  
XX Mycoplasma gallisepticum antigenic protein TTM-1.

DE  
XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1.

XX  
OS Mycoplasma gallisepticum.

XX  
FH Key Location/Qualifiers

FT Misc-difference 262

FT /note= "any amino acid"

FT Misc-difference 283

FT /note= "any amino acid"

XX  
PN JP07133295-A.

XX  
PD 23-MAY-1995.

XX  
PF 27-AUG-1993; 93JP-0213102.

XX  
PR 27-AUG-1993; 93JP-0213102.

XX  
PA (SHIO ) SHIONOGI & CO LTD.

XX  
DR WPI; 1995-220782/29.

XX  
DR N-PSDB; AAQ94711.

XX  
PT A new antigenic protein which reacts with Mycoplasma gallisepticum -  
PT is useful in a component vaccine for use against poultry infected  
PT with M. gallisepticum.

XX  
PS Claim 6; Figs 5-6; 33pp; Japanese.

XX

CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein  
TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum  
CC infectious diseases in poultry, and as a diagnostic agent for  
CC M. gallisepticum.

XX  
SQ Sequence 368 AA;

Query Match 80.3%; Score 1604; DB 16; Length 368;  
Best Local Similarity 94.4%; Pred. No. 4.7e-100;  
Matches 318; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CMSITKDKANPNNGQTOLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60  
|||||  
Db 27 cmsitkdkanpnngqtlqaarmeltdlinakartlasldyakiessaysaeactn 86  
|||||  
Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 120  
|||||  
Db 87 nlnnatleqlkmaaktlnlesainqantdkttfdnehpnlvqpykalkttlegratnleqla 146  
|||||  
Qy 121 STAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTANKNNINNTLSTINEOKTN 180  
|||||  
Db 147 staynqirnnlvlynnasslittktdplnggmildsneittvnrnintlstineqktn 206  
|||||  
Qy 181 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 240  
|||||  
Db 207 adalsnsfikkvionneqsfvgftfnanvqpsnysfvafsdvtpvnykyarrtvxngde 266  
|||||  
Qy 241 PSSRLANTNSITDVSWIYSIAGTNTKYQFSNYGPGSTGYLYFPYKLVKAADANNVGLQ 300  
|||||  
Db 267 pssrlantnsitdvsxiyslagtnkyqfsnygpgstgylyfpyklvkaadannvqlq 326  
|||||  
Qy 301 YKLNGNVQVQVEFATSTANNTTANPTPAVDKIKVAK 337  
|||||  
Db 327 yklngnvqvqvefatstannnttanptqqlmrklk 363  
|||||

## RESULT 8

AAP93646  
ID AAP93646 standard; protein; 235 AA.

XX  
AC AAP93646;

XX  
DT 11-MAY-1990 (first entry)

DE  
XX Amino acid sequence of Mycoplasma gallisepticum (MG1) polypeptide.

XX  
KW Mycoplasma gallisepticum; Poultry vaccine; ss;

XX  
OS Mycoplasma gallisepticum.

XX  
PN EP345021-A.

XX  
PD 06-DEC-1989.

XX  
PF 31-MAY-1989; 89EP-0005441.

XX  
PR 02-JUN-1988; 88JP-0136343.

XX  
PA (JAP ) NIPPON ZEON KK.

XX  
PA (SHIO ) SHIONOGI SEIYAKU KABUSHIKI KAISHA.

XX  
PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

XX  
DR WPI; 1989-358393/49.

XX  
DR N-PSDB; AAN92568.

XX  
PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
PT vaccines.

XX  
PS Disclosure; Fig.1a; 31pp; English.

XX  
CC This amino acid sequence of MG1 is encoded by M1 DNA and elicits an

CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a  
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro  
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.  
 XX  
 SQ Sequence 235 AA;

Query Match 57.2%; Score 1142; DB 10; Length 235;  
 Best Local Similarity 97.0%; Pred. No. 2.9e-69;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 60  
 DB 1 cmsitkkdanpnngqtlgaarmeltdlinakartilasldyakiessaysaeatvn 60  
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 DB 61 nlnnatleqlkmaaktulesainqantdkttfdnehpnlvaykalkttleqratnleгла 120  
 QY 121 STAYNQIRNNLVLYNKASLIKTLDPLNGGTLDSNEITTANKNINNTLSTINEQKTN 180  
 DB 121 staynqirnnlvlynassliktldplnggmldsnaittavrnnintlstineqktn 180  
 QY 181 ADALSNFIFKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 235  
 DB 181 adalsnfikvigneqsfvgtftnanvqpsnyfafsadvtpvnykyarritv 235

## RESULT 9

AAR05081  
 ID AAR05081 standard; protein; 235 AA.  
 XX  
 AC AAR05081;  
 XX  
 DT 08-OCT-1990 (first entry)  
 XX  
 DE MG-1 antigen.  
 XX  
 KW Mycoplasma gallisepticum; poultry; vaccine.  
 XX

JP02111795-A.

24-APR-1990.

02-JUN-1989; 89JP-0136343.

02-JUN-1989; 89JP-0136343.

(JAPG ) NIPPON ZEON KK.

(SHIO ) SHIONOGI KK.

WPI; 1990-169109/22.

N-PSDB; AAQ04686.

Diagnostic and vaccine for poultry mycoplasma serum - utilises  
 antigen protein of the disease and recombinant vector  
 incorporated with its coding gene.

Claim 2; Fig 1a; 20pp; Japanese.

DNA encoding the protein can be inserted into an expression vector  
 for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 be ligated to other DNA to produce fusion proteins with an N-terminal  
 bacterial enzyme sequence.  
 See also AAR05081-2 and AAR06437-41.

SQ Sequence 235 AA;

Query Match 57.2%; Score 1142; DB 11; Length 235;  
 Best Local Similarity 97.0%; Pred. No. 2.9e-69;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 60  
 DB 1 cmsitkkdanpnngqtlgaarmeltdlinakartilasldyakiessaysaeatvn 60  
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 DB 61 nlnnatleqlkmaaktulesainqantdkttfdnehpnlvaykalkttleqratnleгла 120  
 QY 121 STAYNQIRNNLVLYNKASLIKTLDPLNGGTLDSNEITTANKNINNTLSTINEQKTN 180  
 DB 121 staynqirnnlvlynassliktldplnggmldsnaittavrnnintlstineqktn 180  
 QY 181 ADALSNFIFKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 235  
 DB 181 adalsnfikvigneqsfvgtftnanvqpsnyfafsadvtpvnykyarritv 235

## RESULT 10

AAR05082  
 ID AAR05082 standard; protein; 261 AA.  
 XX  
 AC AAR05082;  
 XX  
 DT 08-OCT-1990 (first entry)  
 XX  
 DE TMG-1 antigen.

Mycoplasma gallisepticum; poultry; vaccine.

JP02111795-A.

24-APR-1990.

02-JUN-1989; 89JP-0136343.

02-JUN-1989; 89JP-0136343.

(JAPG ) NIPPON ZEON KK.

(SHIO ) SHIONOGI KK.

WPI; 1990-169109/22.

N-PSDB; AAQ04687.

Diagnostic and vaccine for poultry mycoplasma serum - utilises  
 antigen protein of the disease and recombinant vector  
 incorporated with its coding gene.

Claim 2; Fig 2; 20pp; Japanese.

DNA encoding the protein can be inserted into an expression vector  
 for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 be ligated to other DNA to produce fusion proteins with an N-terminal  
 bacterial enzyme sequence.  
 See also AAR05081 and AAR06437-41.

SQ Sequence 261 AA;

Query Match 57.2%; Score 1142; DB 11; Length 261;  
 Best Local Similarity 97.0%; Pred. No. 3.4e-69;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 60  
 DB 27 cmsitkkdanpnngqtlgaarmeltdlinakartilasldyakiessaysaeatvn 86  
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 DB 87 nlnnatleqlkmaaktulesainqantdkttfdnehpnlvaykalkttleqratnleгла 146  
 QY 121 STAYNQIRNNLVLYNKASLIKTLDPLNGGTLDSNEITTANKNINNTLSTINEQKTN 180

Db 147 stayqirnnlvdlynnasslittkldplngmllsneittvnrnintlstineqktn 206  
 QY 181 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 235  
 Db 207 adalsnsfikkvigneqsfvgtftnannvqpsnysfvafsadvtvnykyarrtv 261

RESULT 11  
 AAR79911  
 ID AAR79911 standard; Protein; 261 AA.  
 AC AAR79911;  
 DT 19-JUL-1996 (first entry)  
 DE M.gallisepticum 261 amino acid protein.  
 KW Detection; probe; primer; PCR; amplification; secretion; lung;  
 KW avian chronic respiratory disease; respiratory tract; nasal cavity.  
 OS Mycoplasma gallisepticum.  
 XX JP07236498-A.  
 XX 12-SEP-1995.  
 XX 25-FEB-1994; 94JP-0052764.  
 XX 25-FEB-1994; 94JP-0052764.  
 XX (JAPG ) NIPPON ZEON KK.  
 XX (SHIO ) SHIONOGI & CO LTD.  
 XX WPI: 1995-347462/45.  
 XX N-PSDB; AAT04076.

Detection of Mycoplasma gallisepticum - for the quick detection,  
 i.e. within one day, of avian chronic respiratory diseases  
 Claim 3; Page 10-11; 11pp; Japanese.  
 This is the amino acid sequence of a 261 amino acid protein encoded  
 by a fragment of the Mycoplasma gallisepticum genome. The encoding  
 sequence and the sequence of AAT04075 (encoding a 661 amino acid  
 protein) can be used to detect M.gallisepticum using probes based on  
 nucleotides 1125-1648 and primers based on nucleotides 449-466, the  
 complement of bases 893-919, 1908-1934 and the complement of bases  
 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA  
 encoding this protein. The method using these sequences is faster i.e. is  
 able to detect M.gallisepticum, which causes avian chronic respiratory  
 diseases, within one day, from avian secretions, washings from the lung,  
 respiratory tract, nasal cavity, etc.  
 Sequence 261 AA;

Query Match 57.28; Score 1142; DB 16; Length 261;  
 Best Local Similarity 97.08; Pred. No. 3.4e-69;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CWSITKDDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSAETVN 60  
 Db 27 cmsitkddanpnngqtlqearmeltldlnakartlasldyakyakleaslsayseaeavn 86  
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 Db 87 nlnnatleqlkmaktlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146  
 QY 121 STAYNQIRNNLVDLYNKASSLITTKTLDPLNGTLLDSNEITTANKNNINTLSTINEQKTN 180  
 Db 147 staynqirnnlvdlynnasslittkldplngmllsneittvnrnintlstineqktn 206

QY 181 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 235  
 Db 207 adalsnsfikkvigneqsfvgtftnannvqpsnysfvafsadvtvnykyarrtv 261  
 RESULT 12  
 AAP93959  
 ID AAP93959 standard; Protein; 261 AA.  
 XX AAP93959;  
 AC AAP93959;  
 DT 11-MAY-1990 (first entry)  
 DE Amino acid (AA) sequence of TMG-1 polypeptide.  
 KW Mycoplasma gallisepticum; Poultry vaccine; ss;  
 OS Mycoplasma gallisepticum.  
 XX EP345021-A.  
 XX 06-DEC-1989.  
 XX 31-MAY-1989; 89EP-0005441.  
 XX 02-JUN-1988; 88JP-0136343.  
 XX (JAPG ) NIPPON ZEON KK; (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.  
 XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 WPI: 1989-358393/49.  
 XX N-PSDB; RAN92574.  
 PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
 PT vaccines.  
 PS Disclosure; Fig.2; 31pp; English.

This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the  
 same sequence as that of a polypeptide expressed in Mycoplasma  
 gallisepticum in nature. When the corresponding DNA sequence is inserted  
 into a recombinant vector used to transform a host the antigen protein  
 produced can be used as a vaccine to prevent and diagnose MG infection.  
 Sequence 261 AA;

Query Match 55.9%; Score 1117; DB 10; Length 261;  
 Best Local Similarity 96.2%; Pred. No. 1.6e-67;  
 Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;  
 QY 1 CWSITKDDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSAETVN 60  
 Db 27 cmsitkddanpnngqtlqearmeltldlnakartlasldyakyakleaslsayseaeavn 86  
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 Db 87 nlnnatleqlkmaktlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146  
 QY 121 STAYNQIRNNLVDLYNKASSLITTKTLDPLNGTLLDSNEITTANKNNINTLSTINE-QKT 179  
 Db 147 staynqirnnlvdlynnasslittkldplngmllsneittvnrn-intlstineqkkt 205  
 QY 180 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 235  
 Db 206 nadalsnsfikkvigneqsfvgtftnannvqpsnysfvafsadvtvnykyarrtv 261  
 RESULT 13  
 AAR63226  
 ID AAR63226 standard; Protein; 661 AA.  
 XX







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:46:22 ; Search time 29.99 seconds  
(without alignments)  
1259.190 Million cell updates/sec

Title: US-09-147-052-2\_COPY\_64\_456

Perfect score: 1998  
Sequence: 1 CMSITKKDANPNNGQTLEA.....SSNENNADKIPGYRRPGTEL 393

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	831.5	41.6	702	2 S48754	major surface prote
2	754	37.7	702	2 S48753	major surface prote
3	744	37.2	650	2 S48751	major surface prote
4	743	37.2	649	2 S48752	major surface prote
5	713	35.7	647	2 A49218	hemagglutinin homo
6	575.5	28.8	386	2 S48755	major surface prote
7	189.5	9.5	320	2 S51560	major surface prote
8	175	8.8	6713	2 B89921	hypothetical prote
9	173	8.7	1302	1 JC6009	surface-located me
10	165.5	8.3	661	2 Ac2422	hypothetical prote
11	164.5	8.2	2481	2 D90011	FatB protein [limp
12	163.5	8.2	3890	2 C89921	hypothetical prote
13	159	8.0	933	2 S41539	fibrogen-binding
14	156	7.8	807	2 B1605	hypothetical prote
15	155	7.8	1072	2 A86827	hypothetical prote
16	155	7.8	4688	2 F82885	hypothetical prote
17	154.5	7.7	682	2 S44131	subtilisin-like pr
18	154.5	7.7	1365	2 T30822	impl protein - Myc
19	151.5	7.6	1237	2 D71850	probable outer mem
20	151	7.6	624	2 PC6003	surface membrane p
21	150.5	7.5	1051	2 T18351	impl protein - Myc
22	145	7.3	2399	2 H71879	toxin-like outer m
23	144	7.2	1107	2 AC0976	probable autotrans
24	143.5	7.2	751	2 T40462	ser-lys rich hypot
25	142	7.1	135	2 B49218	hemagglutinin homo
26	142	7.1	2660	2 E85822	probable invasiv z
27	141.5	7.1	1645	2 F96907	phage-related prot
28	140.5	7.0	3194	2 D71917	toxin-like outer m
29	139.5	7.0	1487	2 AG2560	hypothetical prote

30 139.5 7.0 2401 2 T28676 rhoptry protein -  
31 138.5 6.9 1314 1 TNBYR6 transcription regu  
32 137.5 6.9 820 2 T17519 cell surface anti  
33 137.5 6.9 926 2 AE1130 conserved hypothet  
34 137.5 6.9 989 2 D89852 fibrinogen-binding  
35 137.5 6.9 1524 2 S68553 surface layer prot  
36 137.5 6.9 4152 2 T31102 filamentous hemagg  
37 137 6.9 719 2 S55119 hypothetical prote  
38 136.5 6.8 568 2 E97066 membrane associate  
39 136.5 6.8 1073 2 S14032 kinesin-related pr  
40 136.5 6.8 2271 2 F90073 hypothetical prote  
41 136 6.8 5005 2 F82884 hypothetical prote  
42 135.5 6.8 1085 2 T38378 kinesin-like prote  
43 135.5 6.8 1238 2 A64596 hypothetical prote  
44 135 6.8 2269 2 T28677 rhoptry protein -  
45 134.5 6.7 589 2 B97806 hypothetical prote

#### ALIGNMENTS

RESULT 1

S48754

major surface protein (clone pMGAL.4) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48754

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739

A:Accession: S48754

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAR>

A:Cross-references: EMBL:L28424

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 41.6%; Score 831.5; DB 2; Length 702;  
Best Local Similarity 42.0%; Pred. No. 3.2e-38;  
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

Qy 8 DANPNNG-----QTGLEARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVNN 61  
Db 68 NTNPGNGGGTDNAQAOLAAAKKELSDLLATQNSLSTYADYANIONTLTAAVYTTAKSTD 127  
Qy 62 NLNATLEQLKMAKTLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEQRATNLEGLSS 121  
Db 128 NTSATLEQVKSATSTLTQTAIDTAASSKTSFDEKPNELIKAYNALKETLKKWRNLSGLTD 187  
Qy 122 TAYNOIRNNLDLYNKASSLTKTLDPLNGGLLLDSNEITTANKNNNTLSTINEOKTNA 181  
Db 188 SNFAIKTNLTALYSGGDKVTIKTLDPLM-GTAINLSAVSQANTNISNAVKLETWKTA 246  
Qy 182 DALNSFTKKVIONNEQSPVGFFTNANVPNSYVFASFADVTP-----VNYKYARRTV 235  
Db 247 TVLATSFKVEVLVKKNLGTIDT-TNNEQPGNYSFVGISVDVTTGSDNARNPNSFQRKV 305  
Qy 236 WNGD-----EPSRRLANTNSITDYSWISYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVK 290  
Db 306 WTSNTDILSQPOAEGENQOAPDYSWIYNLTGCMGAKYSITFNYYGPGSTGYLYFPYKLVN 365  
Qy 291 AADANNVGLQYKLNNGNVQOQVEFATS-----TSANN-----TTANPPPAVD 331  
Db 366 SSSDSKVALEYKLNESAVKTIIDFSPSQTSPVASDATRENNRSTAAPAOQSSTEINPAPILD 425  
Qy 332 ELKVAKIVLSGLRFQONTIELSVP-TGEGNMKNVAPMIGNIYLSNENNAADK 382  
Db 426 DIKIAKVTLSNLKFGSNTIEFSVPTTAKEGTSKVAPMIGNIYLSNENNAADK 477

Thu Jun 13 10:00:46 2002

us-09-147-052-2\_copy\_64\_456.rpr

Page 2

RESULT 2

major surface protein (clone pmGAL.3) precursor - Mycoplasma gallisepticum

C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48753

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface protein

A:Reference number: S48751; MUID:95010739

A:Accession: S48753

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAR>

A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62417.1; PID:g535690

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 37.7%; Score 754; DB 2; Length 702;

Best Local Similarity 41.6%; Pred. No. 5, 6e-34;

Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 10 NPNNGQT-----QLEARMELTDLINAKAMTLASLODYAKIEASLSAYSEATVNNLNA 65

DB 69 NPNPGNTTPEQQLAAARKTLTDLGENTNVALYADYAKIQSTLSTAYTAKTASNTSA 128

QY 66 TLBOLKMAKTINLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLSTAYN 125

DB 129 TLENRSASTTLOAADIKAANDKRVDSVNOPLVAAYNNLTKLTKSTLSLEGLSENKYG 188

QY 126 QIRNNLVLYNKASSLITKTLPLNG--GTLSDNSBITTANKNINNTLS--TINEQKTN 181

DB 189 GIKNHLKLEDTGSAITAKTLDTSGERTPLEKVN---ANNKIKMAISPESLKKWKGA 245

QY 182 DALSNFIRKVIQNNQSGFVGTFTNANVOPSNYSFVAFSADVTP-----VNYKYARTVW 236

DB 246 DKF-NEFEKNPLSKELKSTSDTAHQEQPANWFAAYSVDLTNSQNLPNWFAQRKVW 304

QY 237 NGD--EPSSRLANTN-SITDVSWIYSLAGTNTKYQFSESNYSGTGYLFFPKVLKAAAD 293

DB 305 TSENQOPGKTALVSSPSVSATDVSWIYSLAGETKTYLTFFEYGPDPNAFLYLPKVLKAAAD 364

QY 294 ANNVGLOKLNNGVQOQVEF-----ATST-----ATST-----SAN 320

DB 365 SSSVALQYSLNKTSSKLINEKPAETVSTNTDQSENEVATTSTTEARSSYKVLVADEAATS 424

QY 321 NTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSSENNA 380

DB 425 NNEMNHTPTVTDINIAKVTLSGLTFGGTIEFSVPEG-----KVAPMIGNMYLTSNSESQ 479

QY 381 DK1 383

DB 480 VK1 482

RESULT 3

major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum

N:Alternate names: major hemagglutinin pmGA

C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48751; A44793

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface protein

A:Reference number: S48751; MUID:95010739

A:Accession: S48751

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-650 <MAR>

Query Match 37.7%; Score 754; DB 2; Length 702;

Best Local Similarity 41.6%; Pred. No. 5, 6e-34;

Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 10 NPNNGQT-----QLEARMELTDLINAKAMTLASLODYAKIEASLSAYSEATVNNN 62

DB 51 NPGDGGMMNAAAQELAAARMGLTTFVDSKAKNGLYVDYKKTONTLTAKYDAAKTVLDN 110

QY 63 LNATLEOLKMAKTINLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLST 122

DB 111 SSTTQNLNEAKTRLETAATAATSKQTFDEQHAELVKVYKELKTLSTNETATLAPYADA 170

QY 123 AYQIRNNLVLYNKASSLITKTLPLNG--GTLSDNSBITTANKNINNTL--STINQOKTN 180

DB 171 QYAGIKMHLGLSDAGKAITTKTLEPEGDP-LTAGAVTMANTKIVEAIRDEVNLNPKKEN 229

QY 181 ADALSNFIRKVIQNNQSGFVGTFTNANVOPSNYSFVAFSADVTP-----NKKYAR 232

DB 230 ATKLADFVFKVQLVKEKIVGVEAHN-KAQPANYSFVGISVDITGTANGOTSIPNNNYAQ 288

QY 233 RTWV-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSESNYSGTGYL 293

DB 289 RTITNGDEP--RSVSNTPVDGOTMAOPLSNVSWIYSLAGTGAKYTLETTYGPSTGYL 346

QY 284 FPKVLKAAADANNVGLYKLNNGVQOQVEFATSTANNTTAN-PTPAVDIEIKVAKIVLSG 342

DB 347 FPKVLVTSQDKLGLLEYKLDA----TEPSAITFGNEQTMNGKTPVNDINVAKVTLAN 402

QY 343 LRFGQNTIELSVPTGEGNNKVPAMIGNIYLSSENNAADKI 383

DB 403 LIFGSKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 438

RESULT 4

major surface protein (clone pmGAL.2) precursor - Mycoplasma gallisepticum

C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48752

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface protein

A:Reference number: S48751; MUID:95010739

A:Accession: S48752

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-649 <MAR>

A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62416.1; PID:g535689

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 37.2%; Score 743; DB 2; Length 649;

Best Local Similarity 42.3%; Pred. No. 2e-33;

Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

Thu Jun 13 10:00:46 2002

us-09-147-052-2\_copy\_64\_456.rpr

Page 2

RESULT 2

major surface protein (clone pmGAL.3) precursor - Mycoplasma gallisepticum

C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48753

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface protein

A:Reference number: S48751; MUID:95010739

A:Accession: S48753

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAR>

A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62417.1; PID:g535690

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 37.7%; Score 754; DB 2; Length 702;

Best Local Similarity 41.6%; Pred. No. 5, 6e-34;

Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 10 NPNNGQT-----QLEARMELTDLINAKAMTLASLODYAKIEASLSAYSEATVNNLNA 65

DB 69 NPNPGNTTPEQQLAAARKTLTDLGENTNVALYADYAKIQSTLSTAYTAKTASNTSA 128

QY 66 TLBOLKMAKTINLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLSTAYN 125

DB 129 TLENRSASTTLOAADIKAANDKRVDSVNOPLVAAYNNLTKLTKSTLSLEGLSENKYG 188

QY 126 QIRNNLVLYNKASSLITKTLPLNG--GTLSDNSBITTANKNINNTLS--TINEQKTN 181

DB 189 GIKNHLKLEDTGSAITAKTLDTSGERTPLEKVN---ANNKIKMAISPESLKKWKGA 245

QY 182 DALSNFIRKVIQNNQSGFVGTFTNANVOPSNYSFVAFSADVTP-----VNYKYARTVW 236

DB 246 DKF-NEFEKNPLSKELKSTSDTAHQEQPANWFAAYSVDLTNSQNLPNWFAQRKVW 304

QY 237 NGD--EPSSRLANTN-SITDVSWIYSLAGTNTKYQFSESNYSGTGYLFFPKVLKAAAD 293

DB 305 TSENQOPGKTALVSSPSVSATDVSWIYSLAGETKTYLTFFEYGPDPNAFLYLPKVLKAAAD 364

QY 294 ANNVGLOKLNNGVQOQVEF-----ATST-----ATST-----SAN 320

DB 365 SSSVALQYSLNKTSSKLINEKPAETVSTNTDQSENEVATTSTTEARSSYKVLVADEAATS 424

QY 321 NTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSSENNA 380

DB 425 NNEMNHTPTVTDINIAKVTLSGLTFGGTIEFSVPEG-----KVAPMIGNMYLTSNSESQ 479

QY 381 DK1 383

DB 480 VK1 482

RESULT 3

major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum

N:Alternate names: major hemagglutinin pmGA

C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48751; A44793

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface protein

A:Reference number: S48751; MUID:95010739

A:Accession: S48751

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-650 <MAR>

Query Match 37.7%; Score 754; DB 2; Length 702;

Best Local Similarity 41.6%; Pred. No. 5, 6e-34;

Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 10 NPNNGQT-----QLEARMELTDLINAKAMTLASLODYAKIEASLSAYSEATVNNLNA 65

DB 69 NPNPGNTTPEQQLAAARKTLTDLGENTNVALYADYAKIQSTLSTAYTAKTASNTSA 128

QY 66 TLBOLKMAKTINLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLSTAYN 125

DB 129 TLENRSASTTLOAADIKAANDKRVDSVNOPLVAAYNNLTKLTKSTLSLEGLSENKYG 188

QY 126 QIRNNLVLYNKASSLITKTLPLNG--GTLSDNSBITTANKNINNTLS--TINEQKTN 181

DB 189 GIKNHLKLEDTGSAITAKTLDTSGERTPLEKVN---ANNKIKMAISPESLKKWKGA 245

QY 182 DALSNFIRKVIQNNQSGFVGTFTNANVOPSNYSFVAFSADVTP-----VNYKYARTVW 236

DB 246 DKF-NEFEKNPLSKELKSTSDTAHQEQPANWFAAYSVDLTNSQNLPNWFAQRKVW 304

QY 237 NGD--EPSSRLANTN-SITDVSWIYSLAGTNTKYQFSESNYSGTGYLFFPKVLKAAAD 293

DB 305 TSENQOPGKTALVSSPSVSATDVSWIYSLAGETKTYLTFFEYGPDPNAFLYLPKVLKAAAD 364

QY 294 ANNVGLOKLNNGVQOQVEF-----ATST-----ATST-----SAN 320

DB 365 SSSVALQYSLNKTSSKLINEKPAETVSTNTDQSENEVATTSTTEARSSYKVLVADEAATS 424

QY 321 NTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSSENNA 380

DB 425 NNEMNHTPTVTDINIAKVTLSGLTFGGTIEFSVPEG-----KVAPMIGNMYLTSNSESQ 479

QY 381 DK1 383

DB 480 VK1 482

RESULT 3

major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum

N:Alternate names: major hemagglutinin pmGA

C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48751; A44793

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface protein

A:Reference number: S48751; MUID:95010739

A:Accession: S48751

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-650 <MAR>

Query Match 37.7%; Score 754; DB 2; Length 702;

Best Local Similarity 41.6%; Pred. No. 5, 6e-34;

Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 8 DANPNNGO-----TOLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEVTN 60  
DB 48 DTNPGDGGGMNNAASQELAAARMGLTTFDSKAKNGLGVYDYKKTQNTLTTRAYDAKTVL 107  
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRATNLEGLS 120  
DB 108 DNSSSTTQNLNEAKTRLETAIRTAATSQTDFDEQHAELVKVYKELKTTLSNETATLAPYA 167  
QY 121 STAYNOIRNNLVLYNKASSLITKTLDPNLGTLSDSNEITTANKNINNTL--STINQOK 178  
DB 168 DAQYAGIKMHLGSLYDAGKAITTKLEPVEGDP-LTASAVMANTKIYEAIKDEVLPQK 226  
QY 179 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NKYK 230  
DB 227 ENATKLADSFVKQVLVKEKITGVEEAAH-KAQPANYSFVGSDVTGTTGTGTSIPNWDY 285  
QY 231 ARRTVW-NGDEPSSRILANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 281  
DB 286 AORTTFTNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTGAKYKLEFTYYPSTGY 343  
QY 282 LYFPYKLVKAADANNVGLQYKLNNGNVQOFEFATSTSA-----NNTTANPTPAVDIEKVA 336  
DB 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395  
QY 337 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 383  
DB 396 KVTLANLFGSNKIEFVPA-----EKVSPMIGNMYLSSSPNNWNI 437  
RESULT 5  
A49218  
hemagglutinin homolog pmGAL.2 - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 07-Dec-1999  
C:Accession: A49218  
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.  
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut  
A:Reference number: A49218; MUID:93162830  
A:Accession: A49218  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-647 <MAR>  
A:Cross-references: GB:S55216; NID:g265625; PIDN:AAB25397.1; PID:g265626  
A:Experimental source: S6  
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBIP:125183)  
C:Genetics:  
A:Genetic code: SGC3

Query Match 35.7%; Score 713; DB 2; Length 647;  
Best Local Similarity 41.1%; Pred. No. 8.8e-32;  
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

QY 8 DANPNNGO-----TOLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEVTN 60  
DB 48 DTNPGDGGGMNNAASQELAAARMGLTTFDSKAKNGLGVYDYKKTQNTLTTRAYDAKTVL 107  
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRATNLEGLS 120  
DB 108 DNSSSTTQNLNEAKTRLETAIRTAATSQTDFDEQHAELVKVYKELKTTLSNETATLAPYA 167  
QY 121 STAYNOIRNNLVLYNKASSLITKTLDPNLGTLSDSNEITTANKNINNTL--STINQOK 178  
DB 168 DAQYAGIKMHLGSLYDAGKAITTKLEPVEGDP-LTASAVMANTKIYEAIKDEVLPQK 226  
QY 179 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NKYK 230  
DB 227 ENATKLADSFVKQVLVKEKITGVEEAAH-KAQPANYSFVGSDVTGTTGTGTSIPNWDY 285  
QY 231 ARRTVW-NGDEPSSRILANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 281  
DB 286 AORTTFTNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTGAKYKLEFTYYPSTGY 343  
QY 282 LYFPYKLVKAADANNVGLQYKLNNGNVQOFEFATSTSA-----NNTTANPTPAVDIEKVA 336  
DB 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395  
QY 337 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 383  
DB 396 KVTLANLFGSNKIEFVPA-----EKVSPMIGNMYLSSSPNNWNI 437  
RESULT 7  
S51560  
major surface protein (clone pmGAL.6) - Mycoplasma gallisepticum (fragment)  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: S51560; S48757  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.  
A:Title: The organisation of the multigene family which encodes the major cell surface  
A:Reference number: S48751; MUID:95010739  
A:Accession: S51560  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA

QY 278 STGYLYFPYKLVKAADANNVGLQYKLNNGNVQOFEFATSTSA-----NNTTANPTPAVDIE 332  
DB 338 STGYLYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVND 389  
QY 333 IKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 383  
DB 390 INVAKVTLANLFGSNKIEFVPA-----EKVSPMIGNMYLSSSPNNWNI 435  
RESULT 6  
S48755  
major surface protein (clone pmGAL.5) precursor - Mycoplasma gallisepticum (fragment)  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: S48755  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.  
A:Title: The organisation of the multigene family which encodes the major cell surface  
A:Reference number: S48751; MUID:95010739  
A:Accession: S48755  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-386 <MAR>  
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62419.1; PID:g535692  
A:Note: the sequence of residues 385-386 and the corresponding nucleotide sequence ar  
C:Genetics:  
A:Genetic code: SGC3  
A:Start codon: GTG

Query Match 28.8%; Score 575.5; DB 2; Length 386;  
Best Local Similarity 42.3%; Pred. No. 1.5e-24;  
Matches 138; Conservative 50; Mismatches 113; Indels 25; Gaps 10;

QY 8 DANPNNG-----OTQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEVTN 61  
DB 64 NTNPGCGGMDNSAQQLAAAKKELSDLLATQNSNLSTYADYAKIKNDLTAAYTTAETASQ 123  
QY 62 NUNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRATNLEGLS 121  
DB 124 NQATLEQVKNAASTLQTATNAVNEKVFEDENNSLVTAYTNLKTLEGEENTTLAAND 183  
QY 122 TA-YNOIRNNLVLYNKASSLITKTLDPNL-GTLSDSNEITTANKNINNTL--STINEQ 177  
DB 184 SANYGKIKTHLSLYNQAKTITTSIL--LNDAGQSPKNDVVKINKEITDAINPTLLNQ 241  
QY 178 KTNADALSNSFIKKVIONNEQSFVG---TFTNANVOPSNYSFVAFSADVTPV-----VNY 228  
DB 242 KANADWLATSFTKQVL-NDQALTSGSSETSMOTQPGNYSFVGSDVTGTSNNARPNN 300  
QY 229 KYARTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFNSYGPS-TCYLFPYK 287  
DB 301 NPAQRKVDNTNAPLAQTEQSKLTDVSWIYSLSGMGAKYTTVTFDYIGASNAYLYFPYK 360  
QY 288 LVKAADANNVGLQYKLNNGNVQOFEF 313  
DB 361 LVQTNQD--NVGLQYLVNNTTPKLVNF 384  
RESULT 7  
S51560  
major surface protein (clone pmGAL.6) - Mycoplasma gallisepticum (fragment)  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: S51560; S48757  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.  
A:Title: The organisation of the multigene family which encodes the major cell surface  
A:Reference number: S48751; MUID:95010739  
A:Accession: S51560  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA

A:Residues: 1-320 <MAR>  
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62415.1; PID:g535688  
C:Genetics:  
A:Genetic code: SGC3

Query Match 9.5%; Score 189.5; DB 2; Length 320;  
Best Local Similarity 48.9%; Pred. No. 0.0015;  
Matches 46; Conservative 11; Mismatches 26; Indels 11; Gaps 3;

Qy 291 AADANNVQLKNGVQVEFATSTANNTTTPAVDEIKVAKIVLSGLRFGONTI 350  
Db 41 SADS-----PTNQNSQSQNOAPEASA---MNETPTVDGINVAKVLTLDLRFSGNTI 90  
Qy 351 ELSVP-TGEGNNKVPAMIGNIYLSSNENNADKI 383  
Db 91 ELSVPTDEGTSKVPAMIGNIYITSDNQQRKI 124

RESULT 8  
B89921  
hypothetical protein ebhA [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
A:Accession: B89921  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Gi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: B89921  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-6713 <KUR>  
A:Cross-references: GB:BA000018; PID:g13701233; PIDN:BAB42527.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: ebhA

Query Match 8.8%; Score 175; DB 2; Length 6713;  
Best Local Similarity 22.6%; Pred. No. 0.54;  
Matches 100; Conservative 66; Mismatches 179; Indels 98; Gaps 17;

Qy 8 DANNPQQTQLEARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAPT-----VNNN 62  
Db 492 EAVSNTG-TQLNTAMANLQNGINDKANTLAS-ENYHDADSDKKTAYTQAVTNAENILKN 549  
Qy 63 LNATLQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQKATNLEGLSST 122  
Db 550 SGNLD-----KAAVENALSQVTNAKALGNH-NLEQAKSNANTTTI-----NGLQHLTTA 599  
Qy 123 AYQIRNLDVLYNKAS-SLITKTLPLNG--GTLSDSNEITANKNNITLSTINEQKT 179  
Db 600 QKDLKQVQQAONVAGDVTVKSSANTLNGAMGTLRNSIDQNTATKNGQNYLDATERKT 659  
Qy 180 -----NADALNSFTKKVIONNEQSFVGTFTNANVOPSNYSFVAFS 220  
Db 660 NYNNAVDSANGVINATSNPNMDANAINQIATQVTSTKNALDGHNLTOAQKT-----ATN 714  
Qy 221 ADVTPVNYKYARTVWNGDPPSSRLANTNSITDVSWIYSLA-----GNTKY 268  
Db 715 AIDGATNLNKAQDALKAAQVTSQAQVANVTSIQATNELNTAMGQLQHGIDDENATKQTQ 774  
Qy 269 QFSFNYGPGTGYLYFPYKLVKADA-----NNVGLQYKLN-----NGNVQ 309  
Db 775 KYRDAEQSKTAY-----DQVAAKAILNKQTGSNSDKAAVDRLAQVTSKDALNGDAK 830  
Qy 310 QVEPATSTAN-----NTTANPTTTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNN- 362  
Db 831 LAEAKAARQLNGLNHTNAORTALEGQINQATTV-----DGVNTVTKTNTANTLDGAMNS 885

Qy 363 -----KVAPMIGNIYLSSNEN 378  
Db 886 LQAINDKDQATLRNQNYLDADES 908

RESULT 9  
JC6009  
surface-located membrane protein lmp3 precursor - Mycoplasma hominis  
C:Species: Mycoplasma hominis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
R:Adefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.  
J. Bacteriol. 178, 2775-2784, 1996  
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene sys  
A:Reference number: JC6009; MUID:96213016  
A:Accession: JC6009  
A:Molecule type: DNA  
A:Residues: 1-1302 <LAD>  
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336  
C:Genetics:  
A:Gene: lmp3  
A:Genetic code: SGC3  
C:Superfamily: surface-located membrane protein lmp3; tetratrico peptide repeat homolo  
C:Keywords: duplication; membrane protein  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>  
F:957-992/Domain: tetratrico peptide repeat homology <TT1>  
F:993-1026/Domain: tetratrico peptide repeat homology <TT2>  
F:1089-1120/Domain: tetratrico peptide repeat homology <TT3>  
F:1154-1190/Domain: tetratrico peptide repeat homology <TT4>

Query Match 8.7%; Score 173; DB 1; Length 1302;  
Best Local Similarity 22.8%; Pred. No. 0.079;  
Matches 94; Conservative 67; Mismatches 163; Indels 88; Gaps 18;

Qy 4 ITKDKANPNNGQ-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAPT 59  
Db 864 ITKKTFTFNKDKDKVKELEQTRKIDDEFIN---TNKTPDYSTLSELTSKRDSKNSI 919  
Qy 60 NNNLNTLQKMAKTNLESAINQANTDKTTFDN-----EHPN-----LVEAYKAL 105  
Db 920 TNSNKS--DIETANTLQKALAKANTDKQADNLAARSTKEQLNKSISSANTLLAKLTK 977  
Qy 106 KTLQORATNLEGLSSTAYNQI--RNNLDVLYNKASSL-----ITKTLPLNGTLLDSN 158  
Db 978 DNTIQQAQTELEKEVKQANQAVASNTASQSAKSSLDKAVTEITKKTFTFNKDKDKVPR 1037  
Qy 159 EITANKNNITLSTINEQKTNADLSNFIKKVIONNEQSFVGTFTNANVOPSNYSFVA 218  
Db 1038 ELQTRKIDDEFINT-NKTPNYSYTLSELTSK--RDSKNSITNSNKSNDIETANTEL-- 1092  
Qy 219 FSADVTPVNYKYARTVWNGDPPSSRLANTNSITDVSWIYSLAGTNTKYQFSFNYGPS 278  
Db 1093 -----KQALAKAN--TDKAQADNLA-RSTKEQLNKSISSAN 1125  
Qy 279 TGYLYFPYKLVKAAADANNVGLQYKLVNNGVQVEFATSTANNTTANPTTAVDEI--KVA 336  
Db 1126 T-----LAKLTKDNTITQQAQTE--LEKEVKQANQAVASNTVSMQSAKSSLDTKVT 1176  
Qy 337 KIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNI--YLSNENNADKIPGY 386  
Db 1177 EITKKTFTFNKDK-----EAKFNLKKTGRGQIQEFINTNKNKNN-----PNY 1216

RESULT 10  
AG2422  
hypothetical protein all4935 [imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AG2422  
R:kaneko, T.; Nakamura, Y.; Wolke, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AG2422  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-661 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BAB76634.1; PID:g17134073; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all4935

Query Match 8.3%; Score 165.5; DB 2; Length 661;  
Best Local Similarity 22.7%; Pred. No. 0.082;  
Matches 88; Conservative 40; Mismatches 188; Indels 71; Gaps 11;  
Qy 18 LEAARMELTDLI-----NAKAMTLASLDQYAKIEASLSAYSEAEFTVNNNLNATLEQ 69  
Db LKKAEDDLGIIRSTNRKSDANSRLNRAEQDETFQADAGFOTATLNTATNRRNNAQNA 123  
Qy 70 LKMAKTNLESAINQANTDKTTFDNEHFNLYEAYKALKTTLEQRAFNLEGLSSAYNQIRN 129  
Db 124 LNTATNRRNNAQNALNTATNRRDN-----AQNALNTATNRR--NNAQNALNTATNRRN 174  
Qy 130 NLVDLYNKASSLITKTLDPGLNGGTLDSNEITTTANKNNITLSTINEOKTNADALSNSFI 189  
Db 175 NAQNALNTATNRRNNAQNAL-----NTATNRRNNAQNALNTATNRRNNAQAEVDAT 226  
Qy 190 KKVIVQ-----NNRQSPVGTFTNANVQPSYSFVAFSADVTPVNYKYARRTVWNGDEPSSRIL 246  
Db 227 RNLQAARRGNSQKAIQAINALNAQNTALNTATNALTATNTANTAFN-----T 278  
Qy 247 ANTSITDVSIIYSLAGTNTKQFSFGNYGSPSTGYLFFPYKLVKAAANNVGLQYKLNNG 306  
Db 279 ANNFST-----ANTAFNTATNFFSTA-----NTALNTATNFF 311  
Qy 307 NVQOVEATPSTANNTTANPTTAPVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAP 366  
Db 312 NTATLQDQANRLTARN-----DFTANSNFSRTGNELTATNTNPNFNTANTFTATT 365  
Qy 367 MIGNIYLSNNENNADKIPGYRRPGTFL 393  
Db 366 NFNN--ASSRNTAEQARNQVRETRL 390

RESULT 11  
D90011  
FntB protein [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: D90011  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: D90011  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2481 <KUR>  
A:Cross-references: GB:BA000018; PID:g13701961; PIDN:BAB43253.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: fntB(mrp)

Query Match 8.2%; Score 164.5; DB 2; Length 2481;  
Best Local Similarity 23.1%; Pred. No. 0.54;  
Matches 98; Conservative 56; Mismatches 156; Indels 115; Gaps 20;

Qy 8 DANPNNGQTOLE-----AARMELTDLINAKAMTLASLDQYAKIEASLSAYSPAE 57  
Db 1186 DQAKTTGETSIDQVTPPTVKKKATARNETAILNNKLOEIQATPDATDEKQAADA--BAN 1243  
Qy 58 TVNNNLN-----ATLEQLKMAKTNLESAINQANTDKTTFDNEHFNLYEAYKALKTTL-- 109  
Db 1244 TENGKANQAI SAATNAQVDEAKANAIAIN-AVTPKYVKKQAAKDEIDQLQATQNTVIN 1302  
Qy 110 -EQATNLEGLS-----STAYNQIRNL-----VDLYNKASSLITKTLDPGLNGGTL 155  
Db 1303 NQMATNEKEAAIQOLATATVTDKANNITATDONGVDATKAGKNSIQSTQP---ATAV 1359  
Qy 156 DSN-----BITTANKNNITLSTINEOKTNADALSNSFIKKVIONNEQSPVGTFTNAN 208  
Db 1360 KSNKNEVDQAVTTQNAIDNTGATTEKNAAKDL-----VLKAKEAYQDIL--N 1409  
Qy 209 VQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRIILANTNSITDVSIIYSLA----GT 264  
Db 1410 AQTN-----DVTQI-----KQAVADIQGITADTTIKDVAKDELAT 1446  
Qy 265 NTKYQFSFGNYGSPSTGYLFFPYKLVKAAAN-----NVGLQYKLNNGNVQOVEATST 317  
Db 1447 KANEQKAL-----TAQTADATTEKEQANQVDAQLTOGN-QNENASQSI 1490  
Qy 318 SANNTTA-NPTPAVDKIVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAPMIGNIYLSN 376  
Db 1491 DDVNTAKDNAIQADPIQASTDVKTNR-----AELLTEM-----QNKITEILNNNETTNE 1541  
Qy 377 ENNAD 381  
Db 1542 EKGND 1546

## RESULT 12

C89921  
hypothetical protein ebhB [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: C89921  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: C89921  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-3890 <KUR>  
A:Cross-references: GB:BA000018; PID:g13701233; PIDN:BAB42528.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: ebhB

Query Match 8.2%; Score 163.5; DB 2; Length 3890;  
Best Local Similarity 24.4%; Pred. No. 1.1;  
Matches 94; Conservative 54; Mismatches 153; Indels 85; Gaps 20;

Qy 16 TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEA-ETVNNNLNATLEQLKMAK 74  
Db 3528 TDLNTAMGNLQGAINDQETLNS-QNYQDATPSKKTAYTNAVQAARKDILNKSQGNK-TK 3585  
Qy 75 TNLESAINQANTDKTTFDNEHFNLYEAYKALKTTLEQRAFNLEGLSSAYNQIRNLVDL 134  
Db 3586 DQVTEAMNQVNSAKNNLDG-----TRLLDQAKQAKQQLNNMTHTLTAAQKTLNQ---- 3636  
Qy 135 YNKASSLITKTLDPGLNGGTLDSNEITTTANKN-INNTLSTINEQKTNADA--LSNSFI-- 189  
Db 3637 -----INSGTTVAGVHTVOSNANTLQDAMNTLRQSTIANNDATKASEDIYDA 3682  
Qy 190 ---KKVIONNEQSPVGTFTNANVQPS-NYSFVAFSADVTPVNYKYARRTVWNGDE----- 240



```

Db 430 ESNVSDQTSSEAS-TNSNSISLSPNISSTSD--SESATNSSDFSNVAEVANNLASV 486
Qy 64 -NATLEQLKMAKTNLESAINQANTDKTFDNEHPNPLVEAYKALKTTLEQRATN----- 115
Db 487 NNSSSVLSSTSTADNLGINSQSDNLTGDSSEISTGCAFLSSNOTSSEASTNSSSISL 546
Qy 116 -----LEG-LSSTAYNOI-----RNNLDVLYNKASSLI--TKTLDPL-----NGG 152
Db 547 SPSNISTSVLESTTSSSNFENVAEVANNLASVNNSSSVLSSTSTADNLEINQFGSDN 606
Qy 153 TLLDSNEITTANKNINNTLSTINEQKTADALNSFIKKVIONNEQSPVGTFTNANVOPS 212
Db 607 LTKDSSEISTSG-----AFLSSNOTSSEASSNS---MSSINSPSLSLTNSSESAT 655
Qy 213 NVSFVAFSADYTPVNYKYARTVWNGDEPSSRIL-----ANTNSITDVSIIYSLAG 263
Db 656 NQS---NSSEATKVDN-----NSSTHSSNII NSGNSDSDSDSDSSNLSSSPN 704
Qy 264 TNTKYQFSFNYGPFSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQOVEFATSTSANNT- 322
Db 705 LETNOTIS-----SKFEVNNI-----SENPKYVSSNSVQENSTD 740
Qy 323 ---TANPTPAVDE-----TKVAKIVLSGLRFGONTIE---LSVPTGEG 359
Db 741 HEMSTNPKRSISSTSTSSQKESQSNLLNTTEGINNPITFNSSSENSAASILTYSYS 800
Qy 360 NMNKVAPMIGNIYLS-----SNENNADKI 383
Db 801 NNSSESETGCLYISNEAQRONGSEISHLPSSNSNENNVSII 843

```

Search completed: June 12, 2002, 10:49:47  
Job time: 205 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:45:47 ; Search time 22.56 Seconds  
(without alignments)  
425.499 Million cell updates/sec

Title: US-09-147-052-2\_COPY\_64\_456

Perfect score: 1998

Sequence: 1 CWSITKDKANPNNGTQLEA.....SSNENNADKIPGRRPGT 393

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/6C\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	95.6	615	2	US-08-525-742-10
2	1856	92.9	610	2	US-08-525-742-8
3	1638	82.0	368	2	US-08-525-742-4
4	1612	80.7	368	1	US-08-185-851A-4
5	806	40.3	661	2	US-08-525-742-2
6	159	8.0	933	3	US-08-293-728-2
7	159	8.0	933	4	US-09-421-868-2
8	136	6.8	1098	4	US-08-923-992A-8
9	135.5	6.8	1073	4	US-09-541-782-6
10	132.5	6.6	1002	4	US-09-268-347-24
11	131.5	6.6	1164	4	US-08-923-992A-2
12	131	6.6	1004	4	US-09-268-347-30
13	131	6.6	1104	4	US-08-923-992A-4
14	129	6.5	2048	4	US-09-268-347-48
15	126.5	6.3	1128	4	US-08-923-992A-6
16	126.5	6.3	1164	4	US-08-923-992A-10
17	125.5	6.3	518	3	US-09-043-123-2
18	125.5	6.3	3788	4	US-09-336-447A-76
19	124.5	6.2	2411	4	US-09-268-347-36
20	122.5	6.1	1536	1	US-08-038-682-2
21	122.5	6.1	1536	1	US-08-302-832-2
22	122.5	6.1	1536	2	US-08-530-198-2
23	122.5	6.1	1536	2	US-08-469-880-2
24	122.5	6.1	1536	2	US-08-728-470-2
25	122.5	6.1	1536	2	US-08-617-697-2
26	122.5	6.1	1536	4	US-08-719-641-2
27	122.5	6.1	1565	6	5352450-2

28	120.5	6.0	746	5	PCT-US95-10509-2
29	120.5	6.0	1104	4	US-09-268-347-28
30	120.5	6.0	1104	4	US-09-268-347-34
31	119.5	6.0	984	1	US-08-242-932-2
32	119.5	6.0	984	1	US-08-714-481-2
33	119.5	6.0	984	5	PCT-US95-06111-2
34	119.5	6.0	2314	4	US-09-268-347-49
35	119	6.0	2354	4	US-09-268-347-47
36	118.5	5.9	866	2	US-08-483-101-4
37	118	5.9	1612	1	US-08-169-927-2
38	117.5	5.9	889	4	US-09-336-447A-15
39	117.5	5.9	1912	1	US-08-409-995-4
40	117.5	5.9	1912	3	US-08-685-467-4
41	117	5.9	2353	4	US-09-377-155-33
42	117	5.9	2353	4	US-08-913-942-4
43	117	5.9	2353	4	US-09-669-974-33
44	116	5.8	941	4	US-09-336-447A-9
45	116	5.8	1338	2	US-08-728-470-9

ALIGNMENTS

RESULT 1

US-08-525-742-10

; Sequence 10, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

; APPLICANT: Saito, Shuji

; APPLICANT: Ohkawa, Setsuko

; APPLICANT: Saeki, Sakiko

; APPLICANT: Ohsawa, Ikuroh

; APPLICANT: Funato, Hiroo

; APPLICANT: Iritani, Yoshikazu

; APPLICANT: Aoyama, Shigemi

; APPLICANT: Takahashi, Kiyoohito

; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

; TITLE OF INVENTION: AS USE THEREOF

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

; STREET: 1725 K Street, Suite 1000

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,742

; FILING DATE: 25-SEP-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-074139

; FILING DATE: 31-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-245625

; FILING DATE: 30-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP94/00541

; FILING DATE: 31-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Mclelland, Le-Nhung

; REGISTRATION NUMBER: 31,541

; REFERENCE/DOCKET NUMBER: 950811

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-659-2930



APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemitsu  
APPLICANT: Takahashi, Kiyoohito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESSEE: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mclelland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 368 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-4

Query Match 82.08; Score 1638; DB 2; Length 368;  
Best Local Similarity 95.5%; Pred. No. 1.5e-112;  
Matches 322; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
Qy 1 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 60  
Db 27 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86  
Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
Qy 121 STAYNQIRNNLVLYNKASSLITKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180  
Db 147 STAYNQIRNNLVLYNNASSLITKTLDPLNGGMLLDSNEITTVNRNINNTLSTINEQKTN 206  
Qy 181 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 240  
Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 266  
Qy 241 PSSRLANTNSITDVSWIYSIAGTNTKYQSFNSYGPSTGYLYPPYKLVKAADANNVGLQ 300  
Db 267 PSSRLANTNSITDVSWIYSIAGTNTKYQSFNSYGPSTGYLYPPYKLVKAADANNVGLQ 326

Qy 301 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAK 337  
Db 327 YKLNNGNVQVEFATSTSANNTTANPTQOLMRKLK 363  
RESULT 4  
US-08-185-851A-4  
Sequence 4, Application US/08185851A  
Patent No. 5489430  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Fujisawa, Ayumi  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemitsu  
TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene  
TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As  
TITLE OF INVENTION: Well As Vaccines Utilizing the Same  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &  
ADDRESSEE: Naughton  
STREET: 1725 K Street, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0  
SOFTWARE: ASCII from Word Perfect version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185.851A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Theresa M. Stevens-Smith  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: PO-8-A930918  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-887-0357  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 368 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-185-851A-4

Query Match 80.7%; Score 1612; DB 1; Length 368;  
Best Local Similarity 95.0%; Pred. No. 1.2e-110;  
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;  
Qy 1 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 60  
Db 27 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86  
Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
Qy 121 STAYNQIRNNLVLYNKASSLITKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180  
Db 147 STAYNQIRNNLVLYNNASSLITKTLDPLNGGMLLDSNEITTVNRNINNTLSTINEQKTN 206  
Qy 181 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 240  
Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 266

QY 241 PSSRIILANTNSIDYSWIYSLAGTNTKKYQFSPNSYGPSTGYLYFFPYKLVKAADANNVGLQ 300  
 Db 267 PSSRIILANTNSIDYSWIYSLAGTNTKKYQFSPNSYGPSTGYLYFFPYKLVKAADANNVGLQ 326  
 QY 301 YKLNNGNQVEFAFSTSANNTTANPTPAVDEIKVAK 337  
 Db 327 YKLNNGNQVEFAFSTSANNTTANPTQOOLMLKLLK 363

RESULT 5  
 US-08-525-742-2  
 ; Sequence 2, Application US/08525742  
 ; Patent No. 5871742  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saito, Shuji  
 ; APPLICANT: Ohkawa, Setsuko  
 ; APPLICANT: Saeki, Sakiko  
 ; APPLICANT: Ohsawa, Ikuroh  
 ; APPLICANT: Funato, Hirono  
 ; APPLICANT: Iritani, Yoshikazu  
 ; APPLICANT: Aoyama, Shigemitsu  
 ; APPLICANT: Takanashi, Kiyoo  
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
 ; TITLE OF INVENTION: AS USE THEREOF  
 ; NUMBER OF SEQUENCES: 51  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
 ; ADDRESSEE: NAUGHTON  
 ; STREET: 1725 K Street, Suite 1000  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/525,742  
 FILING DATE: 25-SEP-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 05-074139  
 FILING DATE: 31-MAR-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 05-245625  
 FILING DATE: 30-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/JP94/00541  
 FILING DATE: 31-MAR-1994

ATTORNEY/AGENT INFORMATION:  
 NAME: Mclelland, Le-Nhung  
 REGISTRATION NUMBER: 31,541  
 REFERENCE/DOCKET NUMBER: 950811  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-659-2930  
 TELEFAX: 202-8870357  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 661 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-525-742-2

Query Match 40.3%; Score 806; DB 2; Length 661;  
 Best Local Similarity 42.8%; Pred. No. 2,8e-51;  
 Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

QY 8 DANPNNGQTOLEAARMELTD-----LINAKMTLASLDYAKIEASLSAYSEAEV 60  
 Db 45 NTFNSDGGMMNRAAKELADAKAALITTLINGETANLASYEDYAKIKSELTSAYETAKAVS 104  
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 Db 105 AKTGATLNEVNEAKTTLDAALKAASAKNDFDAQHGLSVEAYNNLKETLKEETKLNLSLA 164  
 QY 121 STAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTTANKNNINTTNEOKTN 180  
 Db 165 NENYAAIRTNLSLYEKANTIVTATLDPAT-GNPEVMSVTQANODITNATSRLIANKWN 223  
 QY 181 ADALNSFIKKVIONNEQSFYGTFTNANVQPSNTSFVAFSADVTPVNVKVARRTVMNGDE 240  
 Db 224 ADNLANSFIKQSLVKNNLTRVDVANQOE-QPANYSFVGFVSNNVDTPNWNPFAQRKVAWSEN 282  
 QY 241 -----PSSRIILANTNSIDYSWIYSLAGTNTKKYQFSPNSYGPSTGYLYFFPYKLVKAAD 293  
 Db 283 TPLATTPAEDATQQAASLTDSWYSLNGAEAKYTLTSFRVFGAEKTAIYLYFFPYKLVKTS 342  
 QY 294 ANNYGLOYKLNNNGVQVEF-----ATSTSAN-NTTANP-----TPAVDEIKVAKIV 339  
 Db 343 --NVGLOYKLNGGDKTQINFVQTPASGSDVAANEETMASPAEQMSAPTVDIKIAKVA 400  
 QY 340 LSGURFGONTIELSVPTGEGNMKNKVPIMIGNIYLS-----NENN 379  
 Db 401 LSNLKFSNTEFSVPTG-----KAAPMIGNMYLTSSSEVKNKNIYDDLFGNSFNENN 455  
 QY 380 -----ADKIPY 386  
 Db 456 PTAVTVLLKGY 467

RESULT 6  
 US-08-293-728-2  
 ; Sequence 2, Application US/08293728D  
 ; Patent No. 6008341  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Foster, Timothy J.  
 ; APPLICANT: McDevitt, Damien L.  
 ; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene  
 ; FILE REFERENCE: 05344.105011  
 ; CURRENT APPLICATION NUMBER: US/08/293,728D  
 ; CURRENT FILING DATE: 1994-08-22  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 933  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus aureus  
 US-08-293-728-2

Query Match 8.0%; Score 159; DB 3; Length 933;  
 Best Local Similarity 22.8%; Pred. No. 0.001;  
 Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

QY 2 MSITKKDANPNNGQTOLEAARMELTDLINAKMTLASLDYAKIEASLSAYSEAEV 59  
 Db 73 VSDTKSTNTNGETSVQNPQAQQTQSSSTNATI-----EETPTVGEATT 120  
 QY 60 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 119  
 Db 121 TTNOANTPATQSSNTNABELVNO-TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172  
 QY 120 SST-----AYNOIRNNLVLYNK-----ASSLITKTLPLNGGTL 155  
 Db 173 SITQDPTSTATPSNNEAPQSDASKNKDVVNOAVNTSAPMRAFSLAAVAADAPAGTDI 232  
 QY 156 DSNEITTTANKNNINTTNEQ-----KTN-ADALNSFIK-----KVIONNEQSFYGTFTN 206  
 Db 233 -TNLTNVTGIDSGTIVYPHQAGYVKLYGFSVPNSAVKGDFTKITVPKELNNGVTST 291

QY 207 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRIANTNSITDVSIIYSLAG-TN 265  
Db 292 AKVPP-----IMAGDQ-----VLANGVIDSGNVIYTFDYN 324  
QY 266 TKYQFSFSGYGLYFPYKLVKAADANNVQYKLVNNGVQVQVQVQVQVQVQVQVQV 325  
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362  
QY 326 PTPAVDEIKVAKI-----VLSGLRFQNTIELSVPTGEGNMKNKVPAMI-GNIVLSSNE 377  
Db 363 KTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAVLGNLKPNTDS 422  
QY 378 N 378  
Db 423 N 423  
RESULT 7  
US-09-421-868-2  
; Sequence 2, Application US/09421868  
; Patent No. 6177084  
; GENERAL INFORMATION:  
; APPLICANT: Foster, Timothy J.  
; APPLICANT: McDevitt, Damien L.  
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene  
; FILE REFERENCE: 05344.105011  
; CURRENT APPLICATION NUMBER: US/09/421.868  
; CURRENT FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 08/293,728  
; PRIOR FILING DATE: 1994-08-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 933  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-421-868-2

Query Match 8.0%; Score 159; DB 4; Length 933;  
Best Local Similarity 22.8%; Pred. No. 0.001;  
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;  
QY 2 MSITKKDANPNNGOTQL--EAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATV 59  
Db 73 VSDIKTSNTNGETSVQNPAAQQTQSSNTATT-----EETPVTGEATTT 120  
QY 60 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPLNVEAYKALKTTLEQRTNLEGL 119  
Db 121 TTNOANTPATQSSNTAEELVQ--TSNETTFND--TNTVSSVNS-----PQNSINAENV 172  
QY 120 SST-----AYNQIRNNLVLYNK-----ASSLITKTDPLNGGTL 155  
Db 173 STTQDTSTPATPSNNESQSTDSKNDVQVNAVNTSPRMRASFVAAADAPAGTDI 232  
QY 156 DSNETTTANKNINLTSLINEQ---KTN-ADALSNSFIK---KVIQNNQSFVGTFTN 206  
Db 233 -TNQLTNTVIGDSGTTVYVPHQAGVKLVNNGVSPNSAVKGDGTFKITVPKELNMGVST 291  
QY 207 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRIANTNSITDVSIIYSLAG-TN 265  
Db 292 AKVPP-----IMAGDQ-----VLANGVIDSGNVIYTFDYN 324  
QY 266 TKYQFSFSGYGLYFPYKLVKAADANNVQYKLVNNGVQVQVQVQVQVQVQVQVQV 325  
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362  
QY 326 PTPAVDEIKVAKI-----VLSGLRFQNTIELSVPTGEGNMKNKVPAMI-GNIVLSSNE 377  
Db 363 KTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAVLGNLKPNTDS 422  
QY 378 N 378

Db 423 N 423  
RESULT 8  
US-08-923-992A-8  
; Sequence 8, Application US/08923992A  
; Patent No. 6280738  
; GENERAL INFORMATION:  
; APPLICANT: Tai, Joseph Y.  
; APPLICANT: Blake, Milan S.  
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
; TITLE OF INVENTION: Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923.992A  
; FILING DATE: 05-SEP-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/024,707  
; FILING DATE: 06-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1098 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-923-992A-8

Query Match 6.8%; Score 136; DB 4; Length 1098;  
Best Local Similarity 22.7%; Pred. No. 0.063;  
Matches 84; Conservative 51; Mismatches 159; Indels 76; Gaps 14;  
QY 4 ITTKDANPNNGOTQL--EAARMELTDLINAKAMTLASLDYAKIEASLSAYSE-----AE 57  
Db 192 IRKQAQPDKKEDAEVQVREELGKLFSTKAGLDQIEQHVKKETSSSEENTQKVDEHYAN 251  
QY 58 TVNNNLNATLQLKMAKTNLESAINQANTDKTTFDNEHPLN-----VEAYKALK 106  
Db 252 SLQNLAKSLEELDKATTN-----EQATQVKNQFLENAQKLEIQIPLIKETNVKLYKAMS 306  
QY 107 TTEQRTNLEGLSSTAYNQIRNNLVLYNKASSLITKTDPLNGTLDLSNEITANKN 166  
Db 307 ESLEQVEKELKHNSEA-----NLEDLVAKSKEIVREYEGKLNOSKNLP--ELQKLEEE 357  
QY 167 INNTLSTINEQ-----KTNADALSNSFIKKVI---QNNQSFVGTFTNANVQPSNYFVA 218  
Db 358 AHSKLVQVVEDFRKKFKTSEQVTPKKRLKRDLAANNENQOKI-----ELTVSPENI---- 408  
QY 219 PSADVTPVNYKYARTVWNGDEPSSRIANTNSITDVSIIYSLAGTNTKYQFSF----- 273  
Db 409 -----TVYEGEDVKFTVAKSDSKTTLDFDILL-----TKYNPVSVDRIST 449  
QY 274 NVGPSTG---VLYFPYKLVKAADANNVGLQYKLVNNGVQVQVQVQVQVQVQVQVQV 329

```

; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268.347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-24

Query Match
Best Local Similarity 6.6%; Score 132.5; DB 4; Length 1002;
Matches 91; Conservative 58; Mismatches 153; Indels 175; Gaps 21;

QY 13 NGOTOLEAARM---ELTDLINAKAMTLASLODYAKIEASLSAYSE-----AST 58
Db 362 NGTNPVKISNVADGTEDTDAVSEK--QLKALQD-KQVTLASNAYANGSDADGGKATQT 418
QY 59 VNNLNATLEOLKMAKTNLESA-----INQANTDKTTFDN-----EHPN 97
Db 419 LGNDLNFKEKSTDESELLNIKAAGDVTFTPKKGSVQVGDGKATIQDGAKTTTGLVEASE 478
QY 98 LVEAYKAL-----KTTL-----EQATN----- 115
Db 479 LVDSLNLKLGKVGKVGKGTGATDGTHTDVLVKSQDVKTLKAGDNLKVKQEGTFTYVLRD 538
QY 116 -LEGLSSTAYNQIRNNLVLYNKASSLITK---TLDP---LNGGTLDSNETTANKNIN 168
Db 539 ELTGKVSVEFKOTENGA-----NGASTKIYKDGTLTTPANDANGAATDADKIKVASDGI- 593
QY 169 NTLSTINEOKTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNY 228
Db 594 -----SAGNKAVKNV-----SGLKKFGDANFNP-----LTSSADNLTKOY 629
QY 229 KYARRTVWNGDEPS-----SRILANTNSITDVSWIYS---LAGTWKYQSFESNYG 276
Db 630 DNAYKGLTLDERSKQKQPTVADNTAATVGDRLGWLVSADKTTGESKEYSAQVRNAN 689
QY 277 PSTGYLYFPYKLVKAADANNVGLQYKLNNGNYQVQEFATSTANNTANTPTPAVDEIKYA 336
Db 690 E-----VKFKSGNGINVSQKTLONGTRETIFELAKDEN----- 722
QY 337 KIVLSGLRFGQNTIEL---SVPTGECNM-----NKVAPMIGNIYLSSNEN 378
Db 723 -----AIAFGSGSKALRDNTVAIGTGNVNVNAEKSAGFGDPNTIEDKAGSYAFGNDN 774

RESULT 11
US-08-923-992A-2
; Sequence 2, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
```

```

; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268.347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-24

Query Match
Best Local Similarity 6.8%; Score 135.5; DB 4; Length 1073;
Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;

QY 13 NGO-TOLEAARMELTDLINAKA-MTLASL---QDYAKIEASLSAYSEATVNNLNAT 66
Db 611 NGYFTLLDNFNASWELLNTHSNQLLISMTKITEHFQSLDEALQSARSSCAVENSLLDI 670
QY 67 LEOLKMAKTNLESAINQANTDKTTFDNEHPN-----LVEAYKALKTT---LEORATNLEG 118
Db 671 VSELKSKNSLLDALEHSLQDTSMSOKLNGISSELIQLQDKMESYROLVQELRSLYN 730
QY 119 LSSTAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEIT-----TA 163
Db 731 LQHTHEESQKELMGVNRDIDALVKTCTSLNDADIILSDYISDQSKFESKQODLIANI 790
QY 164 NKINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANVQPSN 213
Db 791 GKIVSNFLOQNESLYTKADILHSHLDNNSIRKANEIMNRSEFL---RNA----- 841
QY 214 YSFVAFSADVTPVNYKYARTVWNG---DEPSSRILANTNSITDVSWIYSLA----- 262
Db 842 ----ASQAEIVGANKERIQTVEGSQLDLSKSKAITHNSRMYD---HCLALAESQKQ 893
QY 263 GTWTKYO-----PFSWYGPSTGYLYPPYKLVKAADANNVGL---QYKLNNGNVQ 309
Db 894 GVNLEQVTLDRLLQKVKEHSEDNTEKHHQOL---LDLLESVGNNDNLIDSIKTPHTELQ 950
QY 310 QVE---FATSTANNT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 350
Db 951 KITDHLVKGTTSNHNTHNELLGDSLSCLNETTIEDLSLVKLETTGDTSPSKRELPAFSP 1010
QY 351 -----ELSVPTGECNMKNVAPMIGNIYLSSNENNAADKIPGYRRP 389
Db 1011 WTRDSSLIKETTNNLSDSKKFVREYITTSNQINPEVDV--YDKP 1052

RESULT 10
US-09-268-347-24
; Sequence 24, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
```

Query Match 6.6%; Score 131.5; DB 4; Length 1164;  
Best Local Similarity 22.8%; Pred. No. 0.15;  
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;  
FILING DATE: 05-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,707  
FILING DATE: 06-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1164 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-923-992A-2

Query Match 6.6%; Score 131.5; DB 4; Length 1164;  
Best Local Similarity 22.8%; Pred. No. 0.15;  
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;  
QY 6 KQANPNNGQTOLEA-ARMELTOLINA-KAMTLASLDQYAKIEASLSAYSE-----AET 58  
DB 230 RQAQQAQKDEAEVKVREBELGLFSKAGLDQEQEHVKETSEENTQKVDEHYANS 289  
QY 59 VNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 107  
DB 290 LQNAQKSLLELDKATN-----EQATQVKNQFLNAQKLEIQPLIKETNVKLYKAMSE 344  
QY 108 TLEQATNLEGLSTAYNOIRNVLVDLYNKASSLIPTKLDPLNGGLDLSNEITTANKNI 167  
DB 345 SLEQVEKELHNSA-----NLEDLVAKSEIVREYEGKLNQSKNLP--ELKQLEEEA 395  
QY 168 NNTLSTINEQ-----KTNDALNSFIKVI---QNEQSFVGTFTNANVOPSNYSFVAF 219  
DB 396 HSKLKQVDEFRKFKTSEQVTPKRVKRDLANENNOQKI-----ELIVSPENI----- 445  
QY 220 SADVPVNYKYARTVWNGDEPSRILANTNSTIDVSWIYSLAGTNTKYQSFSS-----N 274  
DB 446 -----TVVEGEDVKTFTAKSDSKTTLDFSDLL-----TKYNPSVSDRISTN 487  
QY 275 YGPSTG---YLFPYKLVKRAADANNVGLQYKLNNGVQVFEATSTANPTTANP-TPAV 330  
DB 488 YKNTDNHKTAEITIKNLKNSQTVTLKAKDDSGNVVKTFTITVQKKEEQVPTPEQ 547  
QY 331 DEIKVAKIV 339  
DB 548 KDSKTEKV 556

RESULT 12  
US-09-268-347-30  
Sequence 30, Application US/09268347  
Patent No. 6335182  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena M.  
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
FILE REFERENCE: 1038-860  
CURRENT APPLICATION NUMBER: US/09/268,347  
CURRENT FILING DATE: 1999-03-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 30  
LENGTH: 1004  
TYPE: PRT  
ORGANISM: Haemophilus influenzae  
US-09-268-347-30

Query Match 6.6%; Score 131; DB 4; Length 1004;  
Best Local Similarity 20.2%; Pred. No. 0.13;  
Matches 97; Conservative 51; Mismatches 152; Indels 180; Gaps 22;  
QY 13 NGOTQLEAARM-----ELTDLINAKAMTLASLDQYAKIEASLSAYSEA-----ET 58  
DB 359 NGTNPVKISNVADGCTENTDAVSFK--QLKALQD--KQVTLASNAYANGSGDADGKGQT 415  
QY 59 VNNLNATLEQLKMAKTNLESAINQANTDKTTF-----DN----- 93  
DB 416 LSNGLN-----FKFKSTGDELLNIKAENDTVTFTPKKGSVQVGDGKATIODGAKTTTGL 470  
QY 94 -EHPNLVEAYKAL-----KTTL-----EQRATN-- 115  
DB 471 VEASELVDSLNLKLGKVGCTGDTGTYDGTHTDTLVKSGDKVTLKAGDNLKVKREGTNT 530  
QY 116 -----LEGSLSTAYNOIRNVLVDLYNKASSLIKT-----TLDPLNGGLDLSNEITTANKN 166  
DB 531 YALKDELTDVKSVEFKDTANGA-----NGASTKITKGLTITPANGAGAAGA----- 577  
QY 167 INNTLSTINEQKTNDALNSFIKVIKQNNQEQSFVGTFTNANVOPSNYSFVAFSADVTPV 226  
DB 578 --NTANTISVTKDGISA--GNKAVKNV-----SGLKKFGDANFDP-----LTSSADNLTK 624  
QY 227 NYKYARTVWNGDEPS-----SRILANTNSITDVSWIYSLAGT-----NTKYQSFSS 273  
DB 625 QYDNAYKGLTNLDEKSKGKQTPVADNTAATVGDRLGLGHWISADTKGELNKEYNAQVR 684  
QY 274 NYGPSTGYLFPYKLVKRAADANNVGLQYKLNNGVQVFEATSTANPTTANPTPAVDEI 333  
DB 685 NAME-----VKFKSGNGINSGKTLDNCTREITPELAKDEN----- 720  
QY 334 KVAKIVLSGLRFGQNTIEL---SVPTCEGNM-----NKVAPMIGNIYLSSEN 378  
DB 721 -----AIAFGSGKALRDNTVAIGTGNVYNAEKSGAFGDPNIYEDKAGSFAFGNDN 772

RESULT 13  
US-08-923-992A-4  
Sequence 4, Application US/08923992A  
Patent No. 6280738  
GENERAL INFORMATION:  
APPLICANT: Tai, Joseph Y.  
APPLICANT: Blake, Milan S.  
TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
TITLE OF INVENTION: Streptococcal Beta Antigens  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/923,992A  
FILING DATE: 05-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,707  
FILING DATE: 06-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1104 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-923-992A-4

Query Match 6.6%; Score 131; DB 4; Length 1104;  
Best Local Similarity 20.2%; Pred. No. 0.15;  
Matches 85; Conservative 50; Mismatches 154; Indels 132; Gaps 17;

QY 6 KIDANPNNGQTQLEA-ARMEIDDLINA-KAMTLASLDQYAKTEASUSAYSE-----APT 58  
DB 199 RQAOQADKKEDAEVKVREELGKFSSTKAGLDQOQIQEHVKKTSSEENTQKVDEHYANS 258  
QY 59 VNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 107  
DB 259 LONLAQKSLLEELDKATTN-----EQATQVKQNFLENAQKLKEIQLIKETNPKLYKAMSE 313  
QY 108 TLEQRATNLEGLSSTAYNOIRNNLDVLYNKASSLITKTLDPNGGTLTLLDSNEITTANKNI 167  
DB 314 SLEQVEKELKHNSA-----NLQDLVAKSKEIYREYEGKLNQSKNLP--ELKQLEEEA 364  
QY 168 NNTLSTINQ-----KTNDALNSFIKVI---ONNEOSFVGTFTTNANQPSNYSFVAF 219  
DB 365 HSKLQOVVEHFKKTSQVTPKRVKRDLAANNNOOKI-----ELTVSPENI----- 414  
QY 220 SADVTVPYKARYRWNGDEPSSRLANTNSITDVSWIYSLAGTNTKYQFS--FSNYGP 277  
DB 415 -----TYEGEDVKFTVAKSDS-----KTTLDLFDLLTKYNP 447  
QY 278 STGYLYFPYKLVKAADANNVGLQYKLNNGNVQOQVEFATSTANNTANPTPAVDKIVAK 337  
DB 448 SV-----SDRISTNYKTN-----NKVAPMIGNIYLSNENNAADKI 383  
QY 338 IVLSGLRFGQN-TIELSVPTGSGNM-----NKVAPMIGNIYLSNENNAADKI 383  
DB 469 ITIKNLKQSQVTVLAKDDSGNVVEKFTTITVQKKEEKQVPKTPQEKHSKTEQNVQOE 528  
QY 384 P 384  
DB 529 P 529

RESULT 14  
US-09-268-347-48  
Sequence 48, Application US/09268347  
Patent No. 6335182  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena M.  
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
FILE REFERENCE: 1038-860  
CURRENT APPLICATION NUMBER: US/09/268,347  
CURRENT FILING DATE: 1999-03-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 48  
LENGTH: 2048  
TYPE: PRT  
ORGANISM: Haemophilus influenzae  
US-09-268-347-48

Query Match 6.5%; Score 129; DB 4; Length 2048;  
Best Local Similarity 21.2%; Pred. No. 0.48;  
Matches 83; Conservative 60; Mismatches 156; Indels 92; Gaps 19;

QY 4 ITKKDN-----PNNQQTQLEAARMELVDLINAKAMTLASLDQYAKIEASLSAYSEAE 57  
DB 288 VTQSDNNRSPAYTPN---TQALDPKFEQATNNTKAGPLSGNSIKRKI-INVAGVNNKTD 343

QY 58 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTILEQRATNLE 117  
DB 344 AVN-----VAQLEAVVWAKERRITFOGD-DNSTVDKIGLDNTL-----TIK 384  
QY 118 GLSSTAYNOIRNNLDVLYNKA--SSLITKTLDPNGGTLTLLDSNEITTANKNINNTLSTIN 175  
DB 385 GGAET--NALTDDNITGVVKEADNSGLKVKLAKTLNLT-----EVTTLNATTIVKVGCS 437  
QY 176 EOKTNADALNSFIKKVIQNNQSFVGTFTTNANQPSNYSFVAFSADVTVPVNYKARYRTV 235  
DB 438 SSTTAEALLSDSL-----TTPQN----- 456  
QY 236 WNGDEPSSRLANTNSI--TDVSWIYSLAGTN--TKYQFSFSGNYGPGSTGYLYFPYKLVKA 291  
DB 457 -TGSOSTKSTVGVGVKFTNNAETAAIGTTRITRDKIGFARDG-DVDEKQAPYLDKKQ 514  
QY 292 ADANNVGLQYKLNNGVQOQVEFATSTANNTANPTPAVDKIVAKIVL--SGLRFGQNTI 350  
DB 515 LKGVSAI--TIDNG-IDAGNKKISNLAKGSSANDAVTIEQLKAAKPTLNAGAGISVTPT 571  
QY 351 ELSVPTGEGNMKNKVPIMIGNIYLSNENNAAD 381  
DB 572 EISVDAKSGNV--TAPTY-NIGVKTTTELNSD 599

RESULT 15  
US-08-923-992A-6  
Sequence 6, Application US/08923992A  
Patent No. 6280738  
GENERAL INFORMATION:  
APPLICANT: Tai, Joseph Y.  
APPLICANT: Blake, Milan S.  
TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/923,992A  
FILING DATE: 05-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,707  
FILING DATE: 06-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1128 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-923-992A-6

Query Match 6.3%; Score 126.5; DB 4; Length 1128;  
Best Local Similarity 22.5%; Pred. No. 0.32;  
Matches 83; Conservative 51; Mismatches 158; Indels 77; Gaps 15;

```
QY 6 KKDANPNNGQTOLEA-ARMEITDLINA-KAMTLASLODYAKIEASLSAYSE-----AET 58
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
Db 194 RKQAQADKKDAEVKVEELGKLFSSTKAGLDQIHEHVKKETSEENTQKVDEHYANS 253
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
QY 59 VNNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 107
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
Db 254 LONLAQKSLEELDATTN-----EQATQVKNOFLENAQKLKEMOPLIKETNVKLYKAMSE 308
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
QY 108 TLEQRATNLEGLSSTAYNQIRNNLVLYNKASSLITKTLDPNGGTLTLLDSNEITTANKNI 167
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
Db 309 SLEQVEKELKHNSA-----NLEDLVAKSKEIVREYEGKLNQSKNLP--ELKOLEEEA 359
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
QY 168 NNTLSTINEQ-----KTNADALSNSFIKKVI---QNNQSFVGTFTNANVQPSNYSFVAF 219
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
Db 360 HSKLAQVVEDFRKKFKTSEQVTPKKRVKRDLAANNQOKI-----ELTVSPENI----- 409
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
QY 220 SADYTPVNYKYARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFS-----N 274
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
Db 410 -----TVYEGEDVKFTVTAKSQSDSKTTLDFSDLL-----TKYNPSVSDRISTN 451
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
QY 275 YCPSTG---YLYFPYKLYKAADANNVGLQYKLNNGNVQOVVEFATSTSANNTTANP-TPAV 330
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
Db 452 YKNTDNDNHKIAEITIKNLKLNESQVTLKAKDDSGNVVEKFTTITVQKKEEQVKTPEQ 511
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
QY 331 DEIKVAKIV 339
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
Db 512 KDSKTEKV 520
:| | : : | | | | : : | | : : | | : : | | : : | | : : | | :
```

Search completed: June 12, 2002, 10:49:08  
Job time: 201 sec

**THIS PAGE BLANK (USPTO)**

Copyright (c) 1993 - 2000 CompuGen Ltd.

in search, using sw model

June 12, 2002, 10:47:42 ; Search time 17.44 Seconds  
(without alignments)

872.522 Million cell

JS-09-147-052-2\_COPY\_64\_456

Perfect score:

Sequence: 1 CMSITKKDANPNNGQTQLEA.....SSNENNADKIPGYRRPGTFL 393

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

-.Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length:	2000000000
Maximum DB seq length:	2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	154.5	7.7	682	1	NISP_LACLA	lactococcus
2	143	7.2	857	1	AR56_CANAL	c arg5, 6 pr
3	138.5	6.9	1314	1	SW1L_YEAST	saccharomyc
4	137	6.9	719	1	YM41_YEAST	Q03213 saccharomyc
5	135.5	6.8	1085	1	CU7J_SCHPO	P24339 schizosacch
6	132	6.6	1271	1	Y338_MYCGE	P47580 mycoplasma
7	131.5	6.6	1164	1	BAG_STRAG	P27951 streptococc
8	129	6.5	1251	1	RPB2_PLAVB	Q00799 plasmodium
9	128.5	6.4	2334	1	WAPA_BACSU	Q07833 bacillus su
10	128.5	6.4	3712	1	LMA_DROME	Q00174 drosophila
11	128	6.4	995	1	YI09_YEAST	P40442 saccharomyc
12	127	6.4	956	1	YEF3_YEAST	P32618 saccharomyc
13	126.5	6.3	444	1	PST1_YEAST	Q12355 saccharomyc
14	126.5	6.3	1630	1	MSPL_PLAFK	P04932 plasmodium
15	126.5	6.3	1639	1	MSPL_PLAFW	P04933 plasmodium
16	125.5	6.3	1858	1	P3K2_DICDI	P54674 dictyosteli
17	124.5	6.2	821	1	LINS_CAEEL	P45970 caenorhabdi
18	124.5	6.2	1010	1	YK1L_CAEEL	P34278 caenorhabdi
19	124.5	6.2	1018	1	FNBA_STAAU	P14738 staphylococ
20	124	6.2	749	1	MAD1_YEAST	P04957 saccharomyc
21	124	6.2	750	1	YK57_YEAST	P34231 saccharomyc
22	124	6.2	823	1	NSPL_YEAST	P14907 saccharomyc
23	124	6.2	866	1	MYSF_SCHJA	Q05870 schistosoma
24	124	6.2	1159	1	N124_SCHPO	Q09904 schizosacch
25	124	6.2	1433	1	CAT8_YEAST	P39113 saccharomyc
26	123.5	6.2	537	1	ARP_PLAFA	P04931 plasmodium
27	123.5	6.2	1902	1	P2P_LACLC	P15293 lactococcus
28	123	6.2	1260	1	ALSL_CANAL	P46590 candida alb
29	123	6.2	1637	1	MRSP_STAAU	P80544 staphylococ
30	123	6.2	1790	1	US01_YEAST	P25386 saccharomyc
31	122.5	6.1	507	1	FLIC_SALON	Q06974 salmonella
32	122	6.1	1117	1	YN96_YEAST	P53753 saccharomyc
33	122	6.1	1672	1	PMPB_CHLNU	Q9PJY2 chlamydia m

```
or send an email to license@isb-sib.ch).
-----
CC EMBL; L11061; AAA25200.1; -.
CC DR EMBL; X76884; CAA54210.1; -.
CC DR HSSP; P29600; IGCI.
CC DR MEROPS; S08.059; -.
CC DR InterPro; IPR001899; Gram_pos_anchor.
CC DR InterPro; IPR000209; Peptidase_S8.
CC DR Pfam; PF00082; Peptidase_S8; 1.
CC DR PRINTS; PR00723; SUBTILISIN.
CC DR PROSITE; PS00136; SUBTILASE_ASP; 1.
CC DR PROSITE; PS00137; SUBTILASE_HIS; 1.
CC DR PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
CC DR PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
CC DR Hydrolase; Serine protease; Cell wall; Signal; Transmembrane; Zymogen.
CC KW SIGNAL 1 22
CC FT PROPEP 23 195
CC FT CHAIN 196 682
CC FT FT
CC FT FT
CC FT DOMAIN 196 651
CC FT TRANSMEM 652 679
CC FT DOMAIN 680 682
CC FT ACT_SITE 259 259
CC FT ACT_SITE 306 306
CC FT ACT_SITE 512 512
CC FT DOMAIN 652 657
CC FT CONFLICT 500 500
CC FT SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;
-----
Query Match 7.7%; Score 154.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.11; Mismatches 130; Indels 91; Gaps 13;
Matches 75; Conservative 61;
-----
QY 57 ETYNNLNATLEQKMAKTNLESAINQANTDKTTFONEHPNLVEAYKALKTKTLEQRATNL 116
DB 38 ELINHNSNAILSTEGSTTDSINLGAQSPAVKSTTRTE-----LDVTGAAKTLLOTSQVOK 93
QY 117 EGLSSTAYNQIRNLVDLNLKASSLTKLPLDPLNGTLLDSNEITANKNNINTLSTINE 176
DB 94 EMKVSQETQVSSE-----FSKRDSVTNKKEAPVSKDELLEQSEVVVSTSIQKN-KILDN 148
QY 177 OKTNADALSNSFTKKVIQNNQEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV- 235
DB 149 KKRAFNVTSSPLKEKPSNKSADGVINDS-----ASPLSYKAKEVVS 193
QY 236 -----WNGDEPSSRLANTNSITDVSITYSIAGTNTKYQFSNYPGS 278
DB 194 LRQPLKNQKVEAQPILLISNSSEKSKASVYTNSHDFWDYQW-----DMKY---VTNNGES 243
QY 279 TGYLYPPYKLVKADANNVGLQKLVNNGVQVQEFATSTISAN-----NT 322
DB 244 YA-LYQPSKKI-----SVGI-----IDSGIMEHPDLNSLGNFYKNLVPKGGFDNEEPDE 293
QY 323 TANTPTAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKNKVPMTG-NIYLSNEN 378
DB 294 TGNPDSIVD-----KMGHGTAVAGQITANGNLGVAPGITVNIYRVFGEN 338
-----
RESULT 2
AR56_CANAL STANDARD; PRT; 857 AA.
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE ARG5.6 protein, mitochondrial precursor [Contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.6) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)].
GN ARG5.6.
```

```
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001;
RX MEDLINE=97195775; PubMed=9043106;
RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
RT Candida albicans.";
RL Microbiology 143:297-302(1997).
CC -!- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC -!- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate = ADP + N-acetyl-L-
CC glutamate 5-phosphate.
CC -!- PATHWAY: SECOND AND THIRD STEPS IN ARGinine BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X98880; CAA67383.1; -.
CC InterPro; IPR00706; AGPR_act_site.
CC InterPro; IPR001048; Aakinas.
CC InterPro; IPR000534; Semialdh_dh.
CC Pfam; PF00696; aakinas; 1.
CC Pfam; PF01118; Semialdehyde_dh; 1.
CC Pfam; PF02774; Semialdehyde_dhc; 1.
CC ProDom; PD003765; AGPR_act_site; 1.
CC PROSITE; PS01224; ARG; 1.
CC Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;
CC Mitochondrion; Multifunctional enzyme; Transit peptide.
CC TRANSIT 1 MITOCHONDRION (POTENTIAL).
CC CHAIN ? ? ACETYLGLUTAMATE KINASE.
CC CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
CC REDUCTASE.
CC ACT_SITE 669 669 BY SIMILARITY.
CC SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C471 CRC64;
-----
Query Match 7.2%; Score 143; DB 1; Length 857;
Best Local Similarity 23.3%; Pred. No. 0.63;
Matches 99; Conservative 58; Mismatches 134; Indels 134; Gaps 25;
-----
QY 45 IEASLSAY-----SEATVNN---NLNATL-----EOLKMAKTNLESAINQANTD 87
DB 208 VERAINGYLPILTSLAETSSGQLLNVDVAAGELAREFEPLKIVYLNKGGIINGTG 267
QY 88 K-----TTFDNEHNLV-EAYKALKTKLEQR-----ATNLEGLSSTAYNQIRNLVDLYNK 137
DB 268 EKVSAINLDEEYEDLLKESWVYGTUKLKEIHDLQHLPRSSVA-----IIDVNDL 320
QY 138 ASSLITKTLDPLNGTLLDSNEITANKNNIN-NFTLSTINEQNTKADALSNSFKIKQINN 196
DB 321 QKELFTDS-----GAGTL-----IRRGYRLNRLNSLRDFG---NPDLRLNALLR----- 361
QY 197 EQSFVGFTTANVQPSNYSFVAFSADVTPVNYKYARTVNGVDPSSRLANT----- 249
DB 362 -----DPEIKTKGVSVASVYKFLDSVQFKS-----YGDPEP-LEVLAIVVEQNDKI 405
QY 250 -----NSITD-----VSMIYSLAGTNTKYQFSNYSNG-PSTGYL 282
DB 406 PKLDFLSSKTGVLNVDNIFNAIKKDYQKQVNVNENDANLFWYFSKDGSAKNGQI 465
```



QY 39 -LQDYAKIEA-----SLSSAYSEAEVNNLNA-----TLEQLKMAKTNLESAL 81  
 Db 97 LRTSAIAISAPTQTSOPTTEKIGKLSERNVSNVSASNSTTAGTGRMLSQLTN-DSPS 155  
 QY 82 NOANTDK-----TTFDNEHENLV---EAYKALKTTLPQRATNLEGLSTAY----- 124  
 Db 156 NEISTDQLKIFQMDMSARMEMESFNKLSKIAEQTWVNLNKQDNYKVMKLNILL 215  
 QY 125 -----NOIRNL-VOLYNKASSLTKTKLDPL--NGTLLDSNEITANKN--- 166  
 Db 216 KLVAPQARPSTNNAQKLAIELLSISAVSAYLQKMQNNGSGRGHTADLCGDSNTHS 275  
 QY 167 -----INTLSTINEQKNA-----DALNSNFIKKVII 193  
 Db 276 GINQHRTTNGTIDVNTAQLNQFNFSNALTLPDQOHRNNVSNQINSLPNRLGPVI 335  
 QY 194 -----QNNQSFV-GTFTANVQPSNYFVAFSAD----VTPVNYKYARTVWNGDEPS- 242  
 Db 336 NQANQSQOVLIHNTHTQOVNRSIPSPNASTDPFKLPNGIKRRRTQSNNAST 395  
 QY 243 -----SRILANTNSITDVSWIYS-----LAGTNKYQF-SFSNYGPSTGY 281  
 Db 396 NDHASAOKPISALSPLTNSHSTTSMNTNNSIHSVTSASSEFHLNLSNFGTTAL 455  
 QY 282 LYFPYKLVKAADANNVGLQYKLVNNGVQVEFATSTSANNTANPTPAVDEIKVAKIVLS 341  
 Db 456 SLPSLALDNASFPNPONVPIINNTQQLSFLINQDSTTS----- 498  
 QY 342 GLRFGONTIELSYPTGEGNNKVPAMIGNIYLSNENNADKIPGYRRPGT 391  
 Db 499 -----EL-LPSCKSGVNT-----NIV-----NNRSTLPSYKPKMT 529

RESULT 5  
 CUT7\_SCHPO STANDARD; PRT; 1085 AA.  
 AC P24339;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 13-JUL-1998 (Rel. 36, Last annotation update)  
 DE Kinesin-like protein cut7.  
 GN CUT7 OR SPAC25G10.07C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 ON NCBI\_TaxID=4896;  
 RX MEDLINE=91015362; PubMed=2145514;  
 RA Hagan J., Yanagida M.;  
 RT "Novel potential mitotic motor protein encoded by the fission yeast cut7+ gene."  
 RL Nature 347:563-566(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA McLean J., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES; THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO FORM A SHORT SPINDLE THAT ELONGATES TOSPAN THE NUCLEUS AT METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.  
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BINC SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X57513; CAA40738.1; -.  
 DR EMBL; Z70691; CAA94636.1; -.  
 DR PIR; S14032; S14032.  
 DR HSP; P17119; 3KAR.  
 DR InterPro; IPR001752; kinesin.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PRO0380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS00067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;  
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.  
 FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).  
 FT DOMAIN 436 604 COILED COIL (POTENTIAL).  
 FT DOMAIN 715 740 COILED COIL (POTENTIAL).  
 FT DOMAIN 897 955 COILED COIL (POTENTIAL).  
 FT NP\_BIND 159 166 ATP (BY SIMILARITY).  
 FT REPEAT 987 998  
 FT REPEAT 999 1010  
 FT MOD\_RES 1011 1011  
 FT CONFLICT 34 61 PHOSPHORYLATION (BY CDC2) (BY SIMILARITY).  
 FT SASNPKRREPTIDTGYDRSDTNSPT -> LRAILGNVSLLLTL (IN REF. 1).  
 FT SEQUENCE 1085 AA; 122133 MW; 5669277875559D58 CRC64;  
 Query Match 6.8%; Score 135.5; DB 1; Length 1085;  
 Best Local Similarity 20.9%; Pred. No. 2.2;  
 Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;

QY 13 NQO-TQLEARMELTDLINAKA-WTLASL----QDYAKIEASLSSAYSEAEVNNLNAT 66  
 Db 623 NGYFTLLDNFASMEELLNTHSNQLLIISMKTITEHQSLDEALQSSCAVPSNLSLDLI 682  
 QY 67 LEOLKMAKTNLESAINQANTDKTTFDNEHPN-----LVEAYKALKTT---LEORATNLEG 118  
 Db 683 VSELKDSKNSLLDALHSLQDISMSQKLGNGISSLELQDKMKESYQLVQLSLN 742  
 QY 119 LSTAYNQIRNNLVLYNKASSLITKTLPLNGTLLDSNEIT-----TA 163  
 Db 743 LQHTHEESOKELMYGVNRDIDALVKTCTSLNDADIILSDYISDKSKFESKODLIANI 802  
 QY 164 NKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANVQPSN 213  
 Db 803 GKIVSNFLOEQNESLYTKADILHSHLNDNSIRKANEIMNRSSEFL---RNA----- 853  
 QY 214 YSFVAFSADVTPVNYKYARTVWNG-----DEPSSRLANTNSITDYSWISLA----- 262  
 Db 854 ----ASQAEIVGANKERIKQTVENGSQLDSSKSKATHNSRSMYD----HCLALAESQK 905  
 QY 263 GINTKYQ-----SFSNYGPGSTGYLYFPYKLVKAADANNVGL--QYKLVNNGVQ 309  
 Db 906 GVNLEYQTLDRLLQKYKHESEDNTEKHHQOL---LDLLESLVGNNDNLDSIKTPHTELQ 962  
 QY 310 QVE---FATSTSANNT-----ANPTPAVDEIKVAKIVLSGLRFGONTI----- 350  
 Db 963 KITDHLVKGTTSLANTHNTNELLGLGDESCLNLETTIEDTSLVKLETTGDTPSKRELPA 1022  
 QY 351 -----ELSVPTGEGNNKVPAMIGNIYLSNENNADKIPGYRRP 389  
 Db 1023 WTRDSSLIKETTLLNLDSDKKFVRETYTSSNQTNPEPDV--YDKP 1064  
 RESULT 6  
 Y338\_MYCGE STANDARD; PRT; 1271 AA.  
 ID Y338\_MYCGE  
 AC P47580;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical lipoprotein MG338 precursor.





containing the lic and cel loci, and creation of a 177 kb contig  
RT covering the gnt-sacxy region.";  
RL Microbiology 142:3113-3123(1996).  
CC -|- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,  
CC MOTILITY, SECRETION OR DIFFERENTIATION.  
CC -|- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED  
CC INTO THE MEDIUM.  
CC -|- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE  
CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE  
CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED  
CC MOTIF REPEATED 31 TIMES.  
CC -|- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME  
CC SIMILARITY TO THE REPEAT IN E. COLI RHS GROUP OF PROTEINS (RHS-A-D).  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L05634; AAA22883.1; -;  
CC EMBL; D31856; BAA06656.1; -;  
CC EMBL; D29985; BAA06260.1; -;  
CC EMBL; D83026; BAA11683.1; -;  
CC EMBL; Z99124; CAB15959.1; -;  
CC PIR; S32920; S32920.  
CC Subtilist; BG10797; wapa.  
CC InterPro; IPR003305; CBD\_6.  
DR Pfam; PF02018; CBD\_6; 1.  
DR Cell wall; Repeat; Signal; Complete proteome.  
KW SIGNAL 1 28  
FT CHAIN 29 2334  
FT DOMAIN 504 869  
FT REPEAT 504 605  
FT REPEAT 636 736  
FT REPEAT 769 869  
FT DOMAIN 1021 2139  
FT REPEAT 1021 1040  
FT REPEAT 1042 1061  
FT REPEAT 1063 1082  
FT REPEAT 1083 1102  
FT REPEAT 1109 1128  
FT REPEAT 1129 1148  
FT REPEAT 1150 1169  
FT REPEAT 1174 1193  
FT REPEAT 1199 1218  
FT REPEAT 1219 1238  
FT REPEAT 1646 1665  
FT REPEAT 1667 1686  
FT REPEAT 1690 1709  
FT REPEAT 1711 1730  
FT REPEAT 1732 1751  
FT REPEAT 1753 1772  
FT REPEAT 1795 1814  
FT REPEAT 1820 1839  
FT REPEAT 1840 1859  
FT REPEAT 1861 1880  
FT REPEAT 1887 1906  
FT REPEAT 1908 1927  
FT REPEAT 1929 1948  
FT REPEAT 1969 1982  
FT REPEAT 1983 2002  
FT REPEAT 2008 2027  
FT REPEAT 2028 2047  
FT REPEAT 2051 2070  
FT REPEAT 2071 2090  
FT REPEAT 2093 2112  
FT REPEAT 2120 2139  
FT SEQUENCE 2334 AA; 258329 MW; B75138CCD278BAA3 CRC64;  
OR 32 (POTENTIAL).  
3 X 101 AA APPROXIMATE TANDEM REPEATS.  
1-1.  
1-2.  
1-3.  
31 X 21 AA APPROXIMATE TANDEM REPEATS OF  
X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).  
2-1.  
2-2.  
2-3.  
2-4.  
2-5.  
2-6.  
2-7.  
2-8.  
2-9.  
2-10.  
2-11.  
2-12.  
2-13.  
2-14.  
2-15.  
2-16.  
2-17.  
2-18.  
2-19.  
2-20.  
2-21.  
2-22.  
2-23.  
2-24. (APPROXIMATE).  
2-25.  
2-26.  
2-27.  
2-28.  
2-29.  
2-30.  
2-31.

Query Match 6.48; Score 128.5; DB 1; Length 2334;  
Best Local Similarity 22.18; Pred. No. 14;  
Matches 91; Conservative 56; Mismatches 152; Indels 113; Gaps 22;  
QY 3 SITKDDANPNNGQTOLEAARM-ELTDLINAKAMTILASLDYAKIEASLSAYSEATVN- 60  
DB 1772 SVANKEQNTTKKRTFDKNRNLTELTDRGSGQWTVPSPSDKLK---TFSWIHQDKGTNQ 1828  
QY 61 ---NNLNATLSQLKMAKT-----NLESAINQANTDKTTFDNEHNLV----- 99  
DB 1829 FTYNKLQMIEMKDSSTSSYDFYDENGWVQTFIT--GNGGGTSFSYDERNLVSSLHIGDKN 1887  
QY 100 -----EAYK----ALKTTLEQRATNLEGLSSTAYNQLRNLDVLYNKASSLITTTLDPL 149  
DB 1888 GGDILTSEYEDVANGNRTTISSAS--GKVOYEGKL-NQLVKETHEDGVIEYTYDGF 1943  
QY 150 NG----GYLLD-SNEITTPANKNNITLSTINEQKTNADALSNSFIKKVIONNEQSFVGTFF 204  
DB 1944 GNRKTVTTIKDGSSTVNASFNIMNQLTKVNDESISYD-----KNGNRSTDGKF 1992  
QY 205 TNANVPSNYSFVAFSADVTP-VNKKYARRVWNGDEPSSRLANTNSITDVSWIYSLAG 263  
DB 1993 TYTWDADNLTAVTKGDKCPATYKY-----DEKGNRIQKTVNGKV-TNYFYDGS 2043  
QY 264 TTKYQFSESN-----YGPS-----TGILYFEPYKLVKAADANNVGLYKLN-NGN 307  
DB 2044 LNVLYETDADNNVTKSYTYGSGQLLSTYENGKKYF-----YHNAHGD 2087  
QY 308 VQVEFATSTSANNTA-----NP--TPAVDEIKVAKIVLSGLRFGONT 349  
DB 2088 I-----IAISDSTGKTVAKYQYDAMGNPTKTEASDEVKONRYRAGYQYDEET 2135

RESULT 10  
LMA\_DROME  
ID LMA\_DROME STANDARD; PRT: 3712 AA.  
AC Q00174;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Laminin alpha chain precursor.  
GN LANA OR LAMA.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93049203; PubMed=1425586;  
RA Kusche-Gullberg M., Garrison K., Mackrell A.J., Fessler L.I.,  
RA Fessler J.H.;  
RT "Laminin A chain: expression during Drosophila development and  
RT genomic sequence.";  
RL EMBO J. 11:4519-4527(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=94038678; PubMed=8223265;  
RA Henchcliffe C., Garcia-Alonso L., Tang J., Goodman C.S.;  
RT "Genetic analysis of laminin A reveals diverse functions during  
RT morphogenesis in Drosophila";  
RL Development 118:325-337(1993).  
RN [3]  
RP SEQUENCE OF 1762-3712 FROM N.A.  
RX MEDLINE=92078147; PubMed=1744083;  
RA Garrison K., Mackrell A.J., Fessler J.H.;  
RT "Drosophila laminin A chain sequence, interspecies comparison, and  
RT domain structure of a major carboxyl portion.";  
RL J. Biol. Chem. 266:22899-22904(1991).  
CC -|- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF

CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
 CC -|- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.  
 CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC  
 CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE  
 CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES  
 CC IN CELL FATE AND PATTERN, MISSHAPEN LEGS AND DEFECTS IN WING  
 CC STRUCTURE.  
 CC -|- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
 CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT  
 CC MEMBRANES (MAJOR COMPONENT).  
 CC -|- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY  
 CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.  
 CC -|- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO  
 CC DEVELOPMENT AT 10-12 HOURS.  
 CC -|- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -|- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.  
 CC -|- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
 CC -|- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.  
 CC -|- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV) IS NOT  
 CC SIMILAR TO LAMININ DOMAIN IV).  
 CC -|- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; M96388; AAA28662.1; -;  
 CC DR EMBL; L07288; AAC31718.1; -;  
 CC DR EMBL; M75882; AAA28661.1; -;  
 CC DR HSSP; P02468; LITTLE.  
 CC DR FlyBase; FBgn0002526; LanA.  
 CC DR InterPro; IPR000561; EGF-like.  
 CC DR InterPro; IPR001886; LamNT.  
 CC DR InterPro; IPR000034; Laminin\_B.  
 CC DR InterPro; IPR002049; Laminin\_EGF.  
 CC DR InterPro; IPR001791; Laminin\_G.  
 CC DR InterPro; IPR001230; Prenyltn.  
 CC DR Pfam; PF00052; laminin\_B; 1.  
 CC DR Pfam; PF00053; laminin\_EGF; 20.  
 CC DR Pfam; PF00054; laminin\_G; 5.  
 CC DR Pfam; PF00055; laminin\_Nterm; 1.  
 CC DR PRINTS; PRO00011; EGFLAMININ.  
 CC DR ProDom; PD002082; LamNT; 1.  
 CC DR ProDom; PD003031; Laminin\_B; 1.  
 CC DR SMART; SM00180; EGF\_Lam; 17.  
 CC DR SMART; SM00001; EGF\_Like; 1.  
 CC DR SMART; SM00281; LamB; 1.  
 CC DR SMART; SM00282; LamG; 5.  
 CC DR SMART; SM00136; LamNT; 1.  
 CC DR PROSITE; PS00022; EGF\_1; 17.  
 CC DR PROSITE; PS01186; EGF\_2; 5.  
 CC DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 19.  
 CC DR PROSITE; PS50025; LAM\_G\_DOMAIN; 5.  
 CC DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 CC Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 CC FT SIGNAL 1 22 POTENTIAL.  
 CC FT CHAIN 23 3712 LAMININ ALPHA CHAIN.  
 CC FT DOMAIN 25 272 LAMININ N-TERMINAL (DOMAIN VI).  
 CC FT DOMAIN 273 332 LAMININ EGF-LIKE 1.  
 CC FT DOMAIN 333 402 LAMININ EGF-LIKE 2.  
 CC FT DOMAIN 403 447 LAMININ EGF-LIKE 3.  
 CC FT DOMAIN 448 494 LAMININ EGF-LIKE 4.  
 CC FT DOMAIN 495 540 LAMININ EGF-LIKE 5.  
 CC FT DOMAIN 541 586 LAMININ EGF-LIKE 6.

FT DOMAIN 587 631 LAMININ EGF-LIKE 7.  
 FT DOMAIN 632 676 LAMININ EGF-LIKE 8.  
 FT DOMAIN 677 731 LAMININ EGF-LIKE 9.  
 FT DOMAIN 732 784 LAMININ EGF-LIKE 10.  
 FT DOMAIN 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).  
 FT DOMAIN 816 1374 DOMAIN IV'.  
 FT DOMAIN 1375 1420 LAMININ EGF-LIKE 12.  
 FT DOMAIN 1421 1465 LAMININ EGF-LIKE 13.  
 FT DOMAIN 1466 1513 LAMININ EGF-LIKE 14.  
 FT DOMAIN 1514 1564 LAMININ EGF-LIKE 15.  
 FT DOMAIN 1565 1574 LAMININ EGF-LIKE 16 (N-TERMINAL).  
 FT DOMAIN 1575 1775 LAMININ DOMAIN IV (DOMAIN IV).  
 FT DOMAIN 1776 1808 LAMININ EGF-LIKE 16 (C-TERMINAL).  
 FT DOMAIN 1809 1858 LAMININ EGF-LIKE 17.  
 FT DOMAIN 1859 1916 LAMININ EGF-LIKE 18.  
 FT DOMAIN 1917 1969 LAMININ EGF-LIKE 19.  
 FT DOMAIN 1970 2016 LAMININ EGF-LIKE 20.  
 FT DOMAIN 2017 2063 LAMININ EGF-LIKE 21.  
 FT DOMAIN 2064 2111 LAMININ EGF-LIKE 22.  
 FT DOMAIN 2112 2671 DOMAIN II AND I.  
 FT DOMAIN 2672 2868 LAMININ G-LIKE 1.  
 FT DOMAIN 2869 3048 LAMININ G-LIKE 2.  
 FT DOMAIN 3055 3223 LAMININ G-LIKE 3.  
 FT DOMAIN 3249 3528 LAMININ G-LIKE 4.  
 FT DOMAIN 3534 3709 LAMININ G-LIKE 5.  
 FT DOMAIN 3709 3728 COILED COIL (POTENTIAL).  
 FT DOMAIN 3729 3822 COILED COIL (POTENTIAL).  
 FT DISULFID 273 282 COILED COIL (POTENTIAL).  
 FT DISULFID 275 296 BY SIMILARITY.  
 FT DISULFID 298 307 BY SIMILARITY.  
 FT DISULFID 310 330 BY SIMILARITY.  
 FT DISULFID 333 342 BY SIMILARITY.  
 FT DISULFID 335 367 BY SIMILARITY.  
 FT DISULFID 370 379 BY SIMILARITY.  
 FT DISULFID 382 400 BY SIMILARITY.  
 FT DISULFID 403 414 BY SIMILARITY.  
 FT DISULFID 405 421 BY SIMILARITY.  
 FT DISULFID 423 432 BY SIMILARITY.  
 FT DISULFID 435 445 BY SIMILARITY.  
 FT DISULFID 448 460 BY SIMILARITY.  
 FT DISULFID 450 468 BY SIMILARITY.  
 FT DISULFID 470 479 BY SIMILARITY.  
 FT DISULFID 482 492 BY SIMILARITY.  
 FT DISULFID 495 507 BY SIMILARITY.  
 FT DISULFID 497 514 BY SIMILARITY.  
 FT DISULFID 516 525 BY SIMILARITY.  
 FT DISULFID 528 538 BY SIMILARITY.  
 FT DISULFID 541 553 BY SIMILARITY.  
 FT DISULFID 543 560 BY SIMILARITY.  
 FT DISULFID 562 571 BY SIMILARITY.  
 FT DISULFID 574 584 BY SIMILARITY.  
 FT DISULFID 587 599 BY SIMILARITY.  
 FT DISULFID 589 605 BY SIMILARITY.  
 FT DISULFID 607 616 BY SIMILARITY.  
 FT DISULFID 619 629 BY SIMILARITY.  
 FT DISULFID 632 644 BY SIMILARITY.  
 FT DISULFID 634 650 BY SIMILARITY.  
 FT DISULFID 652 661 BY SIMILARITY.  
 FT DISULFID 664 674 BY SIMILARITY.  
 FT DISULFID 677 691 BY SIMILARITY.  
 FT DISULFID 679 700 BY SIMILARITY.  
 FT DISULFID 702 711 BY SIMILARITY.  
 FT DISULFID 714 729 BY SIMILARITY.  
 FT DISULFID 732 746 BY SIMILARITY.  
 FT DISULFID 734 753 BY SIMILARITY.  
 FT DISULFID 755 764 BY SIMILARITY.  
 FT DISULFID 767 782 BY SIMILARITY.  
 FT DISULFID 1375 1387 BY SIMILARITY.  
 FT DISULFID 1377 1394 BY SIMILARITY.  
 FT DISULFID 1396 1405 BY SIMILARITY.

```

FT DISULFID 1408 1418 BY SIMILARITY.
FT DISULFID 1421 1429 BY SIMILARITY.
FT DISULFID 1423 1436 BY SIMILARITY.
FT DISULFID 1438 1447 BY SIMILARITY.
FT DISULFID 1450 1463 BY SIMILARITY.
FT DISULFID 1466 1480 BY SIMILARITY.
FT DISULFID 1468 1487 BY SIMILARITY.
FT DISULFID 1489 1498 BY SIMILARITY.
FT DISULFID 1501 1511 BY SIMILARITY.
FT DISULFID 1514 1526 BY SIMILARITY.
FT DISULFID 1516 1533 BY SIMILARITY.
FT DISULFID 1535 1544 BY SIMILARITY.
FT DISULFID 1547 1562 BY SIMILARITY.
FT DISULFID 1859 1874 BY SIMILARITY.
FT DISULFID 1861 1885 BY SIMILARITY.
FT DISULFID 1887 1896 BY SIMILARITY.
FT DISULFID 1899 1914 BY SIMILARITY.

Query Match 6.4%; Score 128.5; DB 1; Length 3712;
Best Local Similarity 23.4%; Pred. No. 24;
Matches 91; Conservative 48; Mismatches 155; Indels 95; Gaps 19;

Qy 26 TDLINAKAMTILASQDYAKIEPASISAYSEATVNNLNATLEQKAKTNTLESAINQ-- 83
Db 2502 TDLQARQSLQKQVD--DLEPLRNASAGKVKIKISAVNNATEHQLK-----DINKLIDQLP 2555
Qy 84 ANTKDTTFDNEHPNLVEAYKALKTTLEORATNLEGLSTAYNQIR-----NNLVLDLYNKA 138
Db 2556 AESQDDMKNSANASDALEILKNVLE-----ILEPVSQVTPKELEKAHGINRDLDTNKD 2611
Qy 139 SSLTKTLDPLNGTLLDSNEITANKNNINTLSTINEQKTNADALSNSFIKKIQNNEQ 198
Db 2612 VSQANKOLDVEGVS-----SKLNELAEIDIEQQHVRVGSQSLQGEI----- 2654
Qy 199 SFVGTFTNANQPNYSFVAESADVPYNYKYARTVWNGDEPSSRILANTNSITDYSWI 258
Db 2655 -----ENLKAQVBAARQLANSIKV-GVNFK-----PSTILELKTPEKT----- 2691
Qy 259 YSLAGTNTKYQFSNYPGSPSTGYLYFPYKLVKAADANNVGLQYKLVNNGVQVFEATSTS 318
Db 2692 -KLLATRNLTSTYFTTEPS-GFLLYLGNDNKTAQKNDFVAIVEIVNG-----YPLTIDL 2745
Qy 319 ANNTTANPTPAVDEIKVAKIVLSGL-----RFGON-----TIELSVPTGEGNMKVPAMI 368
Db 2746 GNGP-----ERITSDKYVADGRWYQAVVDRMGPNAKLTIRELPNGD-----VVEHS 2792
Qy 369 GNYLSSNEN--NADK-----IPGYRRPG 390
Db 2793 KSGYLEGSONILHYDKNSRLFVGYY--PG 2819

RESULT 11
YIQ9_YEAST STANDARD; PRT; 995 AA.
AC P40442;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 99.7 kDa protein in Sdi1 5' region precursor.
GN YIL169C OR Y19402.07C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;

```

```

RL Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46921; CAA87023.1; -.
DR SGD; S0001431; YIL169C.
DR InterPro; IPR004089; Chemotaxis_transducer.
DR InterPro; IPR000727; T_SNARE.
KW Hypothetical protein; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 995 HYPOTHETICAL PROTEIN YIL169C.
FT CARBOHYD 28 28 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 664 664 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

Query Match 6.4%; Score 128; DB 1; Length 995;
Best Local Similarity 16.1%; Pred. No. 5;
Matches 73; Conservative 100; Mismatches 152; Indels 128; Gaps 15;

Qy 3 SITKDDANPNNGQTOLEAARMELTDLINAKAMTILASQDYAKIEPASISAYSEATVNNN 62
Db 136 SVSQSSSSASDVSSVSSQSSSSASDVSSVSSQSSASD-----VSSVSSQSSASD 188
Qy 63 LNAYLEQLKMAKTNTLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLST 122
Db 189 VSSVSSQSSSSASDVSSVSSQSSSSASDVSSVSSQSSASD-----VSSVSSQSSASD 248
Qy 123 AYNOIRNLVDLYNKASSLITKTLDPNGTLLDSNEITANKNNINTLSTINEQKTNAD 182
Db 249 GSQSVSS-----ASGSSSSFPQST-----SSASTASGASATSNLSSTSSASAS 293
Qy 183 ALSNSFIKK-----VIONEQSF-----VG 202
Db 294 ATASNLSSSSDGTIYLPPTTISGDLTLFGKVIATGEGVVVAAGAKTLTLDGKYSFADLK 353
Qy 203 TFTNANQPNYSFVAESADVPYNYKYARTVWNGDE---PSSRILANTNSITDYSWI 259
Db 354 VYGDLLYKSKETPGTEFDISGENFDVTGN--FNAESAATSASIYFTSSPNSGDI 411
Qy 260 SLA-GTNTKYQFSNYPGSPSTGYLYFPYKLVKAADANNVGLQYKLVNNGVQVFEATST- 317
Db 412 SLSLSKSKKGEVTFSPYNSGAFSF-----SNAI-----LNGGSVSLQRRDDTE 456
Qy 318 -SANTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKVPAMIGNIYL--- 373
Db 457 GSVNNGEIN-----LDNGSTYIVPEVPVSGKGTVNIIS---GNLYLHYP 496
Qy 374 -----SSNENADKIP--GY 386
Db 497 DTFGTQTVVFKGEGVLAVDPTETNATPIPVVGY 529

RESULT 12
YEF3_YEAST STANDARD; PRT; 956 AA.
AC P32618;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 106.1 kDa protein in GLY1-GAL1 intergenic region.
GN YEL043W OR SVGP-ORF14.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

```

OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,  
Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;  
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,  
Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
Taylor P., Wei Y., Yelton M., Bolstein D., Davis R.W.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP FIBRONECTIN TYPE III DOMAIN.  
RX MEDLINE=97148176; PubMed=8994808;  
RA Bateman A., Chothia C.;  
RT "Fibronectin type III domains in yeast detected by a hidden Markov  
model";  
RL Curr. Biol. 6:1544-1546 (1996).  
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: U18779; AAB64999.1; -  
DR PIR: S30834; S30834.  
DR SGD: S0000769; YEL043W.  
DR InterPro: IPR003961; FN\_III.  
DR Pfam: PF00041; fn3; 1.  
DR SMART: SM00060; FN3; 1.  
KW Hypothetical protein.  
FT DOMAIN 35 125 FIBRONECTIN TYPE-III.  
SQ SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;  
  
Query Match 6.4%; Score 127; DB 1; Length 956;  
Best Local Similarity 19.7%; Pred. No. 5.4;  
Matches 104; Conservative 86; Mismatches 163; Indels 176; Gaps 25;  
  
QY 4 ITKKNPNNGQQLAARWELTDL-----INA-KAMTLASLDQYAKTEASLSLSAYSEA 56  
DB 337 LNESNASVANINKEIESLQNEISKMEESKRLNASKSLITSVNANVENDKPIASGEL 396  
QY 57 ETYNNLN-ATLE-----OLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALK 106  
DB 397 SAVLKKLNDFTLEKNGFLSNAGEEFUSKLNADSSLKMKIKQELSIDQE---LEANKLQ 453  
QY 107 TTLEQRATNLEGLSTAYQIRNNLVLYNKASSLITKTL-DPL--NGGTLDSNEITTA 163  
DB 454 SNLLKKSIALE-----NQF--NENSLNRR--NLATKLMVQPKYKNGDSLAATNSNNSA 502  
QY 164 NKN-----TNVLTSTNEOKTNADAL----- 184  
DB 503 EKNRSGSGSIQLPLSNNSMRSTGSDILSNNNKSN--NSNADSAPPLRLHPVSPSPSNEP 560  
QY 185 ---SNSFIKKVIONNE--QSPVGGFTTAN---VQPSNFSFVA-----FSADVTTPVNYK 229  
DB 561 IQPSSLLSQTQDTRNSMLNSHISNNENKQPSYSHALPTTATANATATATATN-G 619  
QY 230 YARRTVWNG---DEPSSRIILA-----NTNSITDVSWIYSLAGTNTKYQFSFSNYG 276  
DB 620 HSRSLNLTWTAQFAQPSHQYQVSTELDQAFFEYDNANHL--ISGLQNMIVDETDYDPNLSNYS 677

QY 277 -----PSTGYLYPPYKLVKAADANNVGLQYKLANGNQ 309  
DB 678 KGFTTDELNDYWTQQPQVQRSTNESLFTTGTGPMSSYK-----ANPVISPT--SSSHLR 729  
QY 310 QVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRF-----GONTIELS-- 353  
DB 730 QTSNATNTNPHQPQLLAATLNDPSLQSFVRSYSGFSYAPQANSLONNINGNETENISPR 789  
QY 354 -----VPTGEGNMKNKVPIMIGNIYLSNENNADKI 383  
DB 750 ISSDFENLLVPNLSPRLSNDVPIVPGNNTTLTPSHSNILTMNHQPTADNI 838  
  
RESULT 13  
PSTL\_YEAST  
ID PSTL\_YEAST STANDARD; PRT; 444 AA.  
AC Q12355;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Protoplast secreted protein 1 precursor.  
GN PST1 OR YDR055W OR D4214 OR YD9609.09.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96381250; PubMed=8789263;  
RA Brandt P., Ramlow S., Otto B., Bloecker H.;  
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm  
of Saccharomyces cerevisiae chromosome IV";  
RL Yeast 12:85-90 (1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP IDENTIFICATION.  
RX MEDLINE=99251092; PubMed=10234784;  
RA Pardo M., Monteoliva L., Pla J., Sanchez M., Gil C., Nombela C.;  
RT "Two-dimensional analysis of proteins secreted by Saccharomyces  
cerevisiae regenerating protoplasts: a novel approach to study the  
cell wall";  
RL Yeast 15:459-472 (1999).  
RN [4]  
RP GPI-ANCHOR.  
RX MEDLINE=20469049; PubMed=11016834;  
RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;  
RT "Up-regulation of genes encoding glycosylphosphatidylinositol  
(GPI)-attached proteins in response to cell wall damage caused by  
disruption of FKS1 in Saccharomyces cerevisiae";  
RL Mol. Gen. Genet. 264:64-74 (2000).  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND  
CC SECRETED BY REGENERATING PROTOPLASTS.  
CC -!- SIMILARITY: BELONGS TO THE SPS2 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: X84162; CAA58971.1; -  
DR EMBL: Z74351; CAA98873.1; -  
DR EMBL: Z49209; CAA89084.1; -  
DR SGD: S0002462; PST1.  
DR COMPLEYEAST-2DPAGE; Membrane; GPI-anchor; Signal.  
KW Glycoprotein; 1 19 POTENTIAL.

FT CHAIN 20 ? PROTOPLAST SECRETED PROTEIN 1.  
FT PROPEP ? 444 REMOVED IN MATURE FORM (POTENTIAL).  
FT DOMAIN 356 416 SER-RICH.  
FT CARBOHYD 57 76 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 444 AA; 45776 MW; 230F60CACA5921A4 CRC64;  
  
Query Match 6.38; Score 126.5; DB 1; Length 444;  
Best Local Similarity 18.19; Pred. No. 2.3; Indels 105; Gaps 15;  
Matches 73; Conservative 79; Mismatches 146; Indels 105; Gaps 15;  
  
Qy 16 TQLEAARME-LTDLINAKAMTASLQDYAKTEASLSAYSEATVNNLNATLQKMAK 74  
Db 88 TSFAADLESITSDLSQSLTILF-----SASFGSLQSVDSIKLTLTFAISSFT 136  
Qy 75 TNLSAINQANTDKTDENEHPNLVEAYKALKTTLEQRATNLEGLSS--TAYNIRNLV 132  
Db 137 SNIKSANNIYSDTSL-----QSDGFSALKVNVFNVNKKLTSIKSPVETVSDSLQ 190  
Qy 133 DLYNKASSLIKTLDPL--NGTLLDSNEITANKN-----INNTLSTINEQKTA 181  
Db 191 FSNQNGOTKI--TFDDLWANNISLTDVHVSFANLQKINSLSLGINNLSLFTKLT 248  
Qy 182 ----DALNSFIKK-----VIQNN-----EQSFVGFTN 206  
Db 249 IGQTFISVNDYLNLSFSLSTIGALVANNTGLQKIGLDNLTTIGGTLEVGVNFTS 308  
Qy 207 ANVQPSNVFVAFADVTPVNYKARTVWNGDEPSSRLANTNSITDVSIIYSLAGTNT 266  
Db 309 LNL--DSLKSVKGADVESKSNFSCNAL-----KALQKKGIGKESVCKNGASST 358  
Qy 267 KYQFSFNYGPGTGYLPYKLVKAADANNVGLQKLNNGNVQVEFATSTSANNTANP 326  
Db 359 SVKLSSTSKSQSS-----QTTAKVSKSSKAEKKFTSGDIKAAASASSVSSSGASS-- 410  
Qy 327 TPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIG 369  
Db 411 -----SSSKSKGNAIMAPIQO-----TTPLVG 434

RESULT 14  
MSPL\_PLAFK  
ID MSPL\_PLAFK STANDARD; PRT; 1630 AA.  
AC P04932;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)  
DE (PMKSA) (P190).  
GN MSP-1.  
OS Plasmodium falciparum (isolate K1 / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=86136024; PubMed=3004972;  
RA Mackay M., Goman N., Bone N., Hyde J.E., Scaife J., Certa U.,  
RA Stunnenberg H., Bujard H.;  
RT "Polymorphism of the precursor for the major surface antigens of

Plasmodium falciparum merozoites: studies at the genetic level. ";  
EMBO J. 4:3823-3829(1985).  
[2]  
REVISIONS, SEQUENCE FROM N.A.  
Pan W., Tolle R., Bujard H.;  
RA Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
(Potential).  
CC -1- PPM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42  
kDa and 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF  
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X03371; CAA27070.1; --  
DR PUR; A25120; SAZOK1  
DR InterPro; IPR000561; EGF-like.  
DR Pfam; PF00008; EGF; 1.  
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;  
Transmembrane; GPI-anchor.  
KW SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.  
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.  
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.  
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 755 755 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 835 835 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;  
  
Query Match 6.38; Score 126.5; DB 1; Length 1630;  
Best Local Similarity 20.78; Pred. No. 11;  
Matches 86; Conservative 62; Mismatches 155; Indels 113; Gaps 19;  
  
Qy 36 LASLQDYAK-----IEASLSAYSEATVNNLNATLQKMAKTNLESAINQANTDKTT 90  
Db 226 VGMEDYIKKKKKTIENTINELIESKKYIDKKNKATKEEK-----KKLYQAYDLSI 278  
Qy 91 FDNEHPNLVEAYKALKTTLEQRATNLEGLSSATYQIRNNLVLDYNKASSLTITKLDPLN 150  
Db 279 YNKQ---LEEAHN-LISVLEKRIDTLKK-----NENIKELDKINEI--KNPPAN 323  
Qy 151 GG-----TLDSN-EITANKNNINNTLSTINEOKTNADAL-----SNSIK 190  
Db 324 SGNTPTLTDNKKKEEHEKEIKETAKTI---KFNIDSLFTDPLEYLEYLRKKNKIDIS 380  
Qy 191 KVIQNEOSFVGTFTNANVQPSNVSFV-----AFSADVTVPVNYKARTVWNGDPS 242  
Db 381 AKVETKESTEPNEYPNGVTYPLSYNDINNALNELNSFGDLNPFDT-----KPS 431  
Qy 243 SRIAN-----TNSITDVSIIYSLAGTNTKYQFS 271  
Db 432 KNIYDNERKKEINEKIEKKIESDKKSYEDRSKSLNDITKEKLE-LNEIYDSK 490  
Qy 272 FSNYGPSTGY-----LYFPYKLVKAADANNVGLQKLNNGNVQVEFATSTSANNTANP 326  
Db 491 FNNIDLTNFEKMGKRYSYKVEKLTH-HNTFASVENSKHNEKLTALKYMEDYSLRN- 548





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:47:17 ; Search time 48.58 Seconds  
(without alignments)  
1399.486 Million cell updates/sec

Title: US-09-147-052-2\_COPY\_64\_456  
Perfect score: 1998  
Sequence: 1 CMSITKDKANPNNGTQLEA.....SSNENADKIPGYRRPGTEL 393

- Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17299429 residues

• Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organellae:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1200	60.1	271	2	Q49464 mycoplasma
2	837	41.9	671	2	Q918D6 mycoplasma
3	831.5	41.6	702	2	Q49499 mycoplasma
4	826	41.3	632	2	Q9XC88 mycoplasma
5	805	40.3	584	2	Q9KH13 mycoplasma
6	761	38.1	680	2	Q9KH14 mycoplasma
7	754	37.7	702	2	Q49498 mycoplasma
8	746	37.3	645	2	Q9ZHD1 mycoplasma
9	745	37.3	644	2	Q9ZHR9 mycoplasma
10	744	37.2	650	2	Q49495 mycoplasma
11	743	37.2	649	2	Q49497 mycoplasma
12	713	35.7	647	2	Q49468 mycoplasma
13	704	35.2	656	2	Q9KH15 mycoplasma
14	690.5	34.6	703	2	Q05122 mycoplasma
15	619.5	31.0	419	2	Q9L8D5 mycoplasma
16	575.5	28.8	386	2	Q49500 mycoplasma

17	557	27.9	367	2	Q9XCG7	Q9XCG7 mycoplasma
18	189.5	9.5	320	2	Q49496	Q49496 mycoplasma
19	175	8.8	6713	16	Q99054	Q99054 staphylococ
20	173	8.7	6713	16	Q99054	Q99054 staphylococ
21	173	8.7	1302	2	Q49547	Q49547 mycoplasma
22	164.5	8.2	2481	16	Q990R6	Q990R6 staphylococ
23	163.5	8.0	3890	16	Q99053	Q99053 staphylococ
24	160.5	8.0	433	16	Q97TP6	Q97TP6 clostridium
25	160.5	8.0	682	2	Q48674	Q48674 lactococcus
26	159	8.0	933	2	Q53653	Q53653 staphylococ
27	158.5	7.9	1344	2	Q49545	Q49545 mycoplasma
28	156	7.8	807	5	Q96262	Q96262 plasmodium
29	155	7.8	1072	16	Q9CF64	Q9CF64 lactococcus
30	155	7.8	4688	16	Q9PQ08	Q9PQ08 ureaplasma
31	154.5	7.7	1365	2	Q49525	Q49525 mycoplasma
32	153.5	7.7	1795	2	Q9LCH3	Q9LCH3 staphylococ
33	153.5	7.7	2478	2	Q9RL69	Q9RL69 staphylococ
34	151.5	7.6	1237	16	Q9ZK57	Q9ZK57 helicobacte
35	151.5	7.6	2478	2	Q9LCH2	Q9LCH2 staphylococ
36	151	7.6	624	2	Q49548	Q49548 mycoplasma
37	150.5	7.5	604	5	Q26021	Q26021 plasmodium
38	150.5	7.5	1051	2	Q49524	Q49524 mycoplasma
39	149.5	7.5	962	2	Q49546	Q49546 mycoplasma
40	148	7.4	716	2	Q49526	Q49526 mycoplasma
41	147.5	7.4	655	5	Q26109	Q26109 plasmodium
42	145	7.3	2399	16	Q9ZK59	Q9ZK59 helicobacte
43	144	7.2	1107	2	Q9F2D8	Q9F2D8 salmonella
44	143.5	7.2	751	3	Q94358	Q94358 schizosacch
45	143	7.2	1034	3	Q9Y7E9	Q9Y7E9 candida gla

ALIGNMENTS

RESULT	1
Q49464	
ID	Q49464 PRELIMINARY; PRT; 271 AA.
AC	Q49464;
DT	01-NOV-1996 (TREMBLrel. 01, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE	01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE	TM-1 (FRAGMENT).
GN	TM-1.
OS	Mycoplasma gallisepticum.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC	Mycoplasmataceae; Mycoplasma.
OX	NCBI_TaxID=2096;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=94025893; PubMed=8212828;
RA	Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,
RA	Kanogawa K., Aoyama S., Iritani Y., Hayashi Y.;
RT	"Cloning and DNA sequence of a 29 kilodalton polypeptide gene of
RT	Mycoplasma gallisepticum as a possible protective antigen.";
RL	Vaccine 11:1061-1066(1993).
DR	EMBL: S65869; AAB28343.2; -
FT	NON_TER 271 271
SQ	SEQUENCE 271 AA; 29817 MW; 8B25DE0CD5C85CA2 CRC64;

Query Match	60.1%;	Score 1200;	DB 2;	Length 271;
Best Local Similarity	97.1%;	Pred. No. 6.2e-52;		
Matches 238;	Conservative 3;	Mismatches 4;	Indels 0;	Gaps 0;
Qy	1	CMSITKDKANPNNGTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN	60	
Db	27	CMSITKDKANPNNGTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN	86	
Qy	61	NNLNATLQOLKMAKNLSEAINQANTDKTTFDNEHNLVEAYKALKTTLLEQRTNLEGLS	120	
Db	87	NNLNATLQOLKMAKNLSEAINQANTDKTTFDNEHNLVEAYKALKTTLLEQRTNLEGLA	146	
Qy	121	STAYNQIRNNLVLDYNKASSLITKTLDPLNGGTLTLDSSNETTANKNINNTLTSTINQKTN	180	

Db 147 STAYQIRNNLVLDYNNASSLITKTLPLNGGMLLDSNEITVNRNINNTLSTINEQKTN 206  
Qy 181 ADALSNFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 240  
Db 207 ADALSNFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 266  
Qy 241 PSSRI 245  
Db 267 PSSRI 271  
RESULT 2  
Q9L8D6 PRELIMINARY; PRT; 671 AA.  
AC Q9L8D6  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE PMGA-LIKE PROTEIN 9.2.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN SIGNAL  
RP SEQUENCE FROM N.A.  
RC STRAIN=F;  
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,  
RA May J.B., Hughtett M.B.; from the F-strain (vaccine strain) of  
RT "A novel pmga-like gene from the F-strain (vaccine strain) of  
RT Mycoplasma gallisepticum";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF210770; AAF29524.1; .  
SQ SEQUENCE 671 AA; 71898 MW; 4D6A6B59175D679 CRC64;

Query Match 41.9%; Score 837; DB 2; Length 671;  
Best Local Similarity 43.4%; Pred. No. 1.1e-33;  
Matches 179; Conservative 75; Mismatches 118; Indels 40; Gaps 9;  
Qy 8 DANP-----NNGQTOLEARMELTDLINAKAMTTLASLODYAKIESLSAYSEATVNN 61  
Db 55 DTNPGNDGGMENSAQAQAAAKKELSDLLATQSNLAKYADYTNQIOTLTAAYTTAKSTSD 114  
Qy 62 NLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRTATNLEGLSS 121  
Db 115 NTSVTLQVKSATSTLQAADTAASAKTSFDEKNPELIRAYALKETLKNETVLSGLTD 174  
Qy 122 TAYNQIRNNLVLDYNNASSLITKTLPLNGGMLLDSNEITVNRNINNTLSTINEQKTNA 181  
Db 175 SNEATIKTNLTALYQSGKDFVKATLDPVSGNA-PQIADITKADKDIADAVSKLETWKTN 233  
Qy 182 DALSNFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNN 237  
Db 234 NTLATSFVKEVLKKNLTGIDT-TNNREQPGNISFVGSYNATNNNEIPWNAQKRVWT 292  
Qy 238 GDEPSSRILA-----NTNSITDWSIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVK 293  
Db 293 SDNGRTSLIESTSDNSSTLTVESWIYSLGAGTKYSLTFNFGPSTGYLYFPYKLVK 352  
Qy 294 ANNVGLOKLNNGVQVEFA-----TSTSA-----NNTT-----ANPTPAVD 331  
Db 353 ENNVALQYTLNLSGSAQEVNAPTAKTSVSADSSGDSNNOTESAETMPVTSDLNPAPTVS 412  
Qy 332 EIKVAKIVLSLGRGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNAK 383  
Db 413 DINIAKTLNLKFGSNTIEFSVTEPS--NKVAPMIGNIYLSNENNAK 462  
RESULT 3  
Q9499 PRELIMINARY; PRT; 702 AA.  
ID Q9499  
AC Q9499;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE PMGAI.4 PROTEIN PRECURSOR.  
GN PMGAI.4.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN SIGNAL  
RP SEQUENCE FROM N.A.  
RC STRAIN=S6;  
RX MEDLINE=95010739; PubMed=7925999;  
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
RA Browning G.F., Whithear K.G., Walker I.D.;  
RT "The organisation of the multigene family which encodes the major cell  
RT surface protein, pmga, of Mycoplasma gallisepticum";  
RL FEBS Lett. 352:347-352(1994).  
DR EMBL: L28424; AAA62418.1; .  
KW Signal.  
SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;  
Query Match 41.6%; Score 831.5; DB 2; Length 702;  
Best Local Similarity 42.0%; Pred. No. 2.1e-33;  
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;  
Qy 8 DANPNNG-----QTOLEARMELTDLINAKAMTTLASLODYAKIESLSAYSEATVNN 61  
Db 68 NTPNGGGTGNAAQAQAAAKKELSDLLATQSNLSTADYANIQNTLTAAYTTAKSTSD 127  
Qy 62 NLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRTATNLEGLSS 121  
Db 128 NTSATLEQVKSATSTLQTAIDTAASAKTSFDEKNPELIRAYALKETLKNENSLGLTD 187  
Qy 122 TAYNQIRNNLVLDYNNASSLITKTLPLNGGMLLDSNEITVNRNINNTLSTINEQKTNA 181  
Db 188 SNEATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAVSQANTNISNAVSKLETWKTN 246  
Qy 182 DALSNFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTV 235  
Db 247 TVLATSFVKEVLKKNLTGIDT-TNNQEQPGNISFVGSYVDVTTGSDNAPNHSFAQKV 305  
Qy 236 WNGD-----EPSSRILANTNSITDWSIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVK 290  
Db 306 WTSNTDILSOPQPAEENQOQSAQSDVSWIYNLTGAKYKSLTFNFGPSTGYLYFPYKLVN 365  
Qy 291 AADANNVGLQYKLNNGVQVEFATS-----TSANN-----TTANPTPAVD 331  
Db 366 SSQSKVALEYKLNESAVKTIQSPQTSVPASDATRENNRSTAAAPQAGSTEINPAPTLD 425  
Qy 332 EIKVAKIVLSLGRGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNAK 382  
Db 426 DIKIAKTLNLKFGSNTIEFSVPTTAKETGSKVAPMIGNIYLSNENNAK 477  
RESULT 4  
Q9XCG8 PRELIMINARY; PRT; 632 AA.  
ID Q9XCG8  
AC Q9XCG8;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE VLHAI PRECURSOR (FRAGMENT).  
GN VLHAI.  
OS Mycoplasma imitans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=29560;  
RN SIGNAL  
RP SEQUENCE FROM N.A.  
RC STRAIN=4229;

```

RX MEDLINE=99392472; PubMed=10463176;
RA Markham P.F., Duffy M.F., Glew M.D., Browning G.F.;
RT "A gene family in Mycoplasma imitans closely related to the pmga
family of Mycoplasma gallisepticum.";
RL Microbiology 145:2095-2103(1999).
DR EMBL; AF141940; AAD39483.1; -.
KW Signal.
FT NON_TER
FT SIGNAL
FT CHAIN
SQ SEQUENCE 632 AA; 66959 MW; 173F5B12E705BE47 CRC64;

Query Match 41.3%; Score 826; DB 2; Length 632;
Best Local Similarity 45.3%; Pred. No. 3.4e-33;
Matches 183; Conservative 65; Mismatches 120; Indels 36; Gaps 9;

Qy 7 KDANPNNGOT-----QLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAE 57
Db 34 KPNPDTGNTNPGDGTDAKQLADAKASLNTLLGSGSTNVALYEDYAKIKDTLSSAYASQ 93
Qy 58 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEORATNLE 117
Db 94 TTANNANATKDVNDAAKALQTAISDAANAKTEFDKANGLVSAKALKETLTKSETNLD 153
Qy 118 GLSSTAYNOIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITANKNNINTLSTINEQ 177
Db 154 GLSDQNSAIAKANLVSLYNKAKDFTTVTLDP-TSGMPKVDIEITSANTAITQAVSAIDSQ 212
Qy 178 KTNADALNSPSFKKIVQNNQESFGYGTFTN----ANVQPSNYSFVAFSADVTP----- 225
Db 213 KTNADTAATTFEIKELDSAKLT-PGTTAEGQQAASQPGNYSFVGFSDVITGRGSEQ 271
Qy 226 --VNYKYARRVWNGDEPS-----SRILANTNSITDVSWIYSLAGTNTKYQFSNPGSTG 280
Db 272 DLPSWNEAKRWVWTEGLSAGOTLVSESTPLTDVSWIYSLTGAISKYTLTFTYGPSTA 331
Qy 281 YLYFPYKLVKAADANNVGLQYKLANG-NVQOVERATPSANNTTANPTTPAYDEIKVAKIV 339
Db 332 YLYFPYKLVQSSDNNKGLQKLNSETLVPTFGNETN----DSGATPAIDDIKVEKVT 387
Qy 340 LSLGRFGONTIELSVPTGEGNNKVPAMIGNIYLSNNENNADKI 383
Db 388 LSNLNFQDNTIEFSVATEE---NKVAPMIGNMYLTSSNNVDKI 428

RESULT 5
Q9KH13 PRELIMINARY; PRT; 584 AA.
AC Q9KH13;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE ADHESIN PMGA1.4 (FRAGMENT).
GN PMGA1.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR InterPro: IPR001986; EPPS_synase.
DR PROSITE: PS00104; EPPS_SYNTHASE_1; UNKNOWN_1.
FT NON_TER 584
SQ SEQUENCE 584 AA; 62453 MW; 5C467BA5F5B27A72 CRC64;

Query Match 38.1%; Score 761; DB 2; Length 680;
Best Local Similarity 39.8%; Pred. No. 5.7e-30;
Matches 164; Conservative 71; Mismatches 129; Indels 48; Gaps 7;

Qy 8 DANPNNGOTQLEAARME-----LTDLINAKAMTSLASLDYAKIEASLSAYSAEYVN 60
Db 63 DTNPGGGQGMNATNQELVNAKKAALSDLIGESKTVELYADYAKIKADLTSAVAKTTIS 122
Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEORATNLEGLS 120
Db 123 DSSTSTLDQVKATATSLQTAINTAASDKKFDQNSQLLMAYKVLKLDTLNKKAEIVMSLN 182
```

```

Query Match 40.3%; Score 805; DB 2; Length 584;
Best Local Similarity 42.4%; Pred. No. 3.4e-32;
Matches 180; Conservative 72; Mismatches 113; Indels 60; Gaps 10;

Qy 10 NPNNGOT----QLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAEYVN 65
Db 69 NPNNGTTPTEQQLAAARKTLTDLTGENTNVALYADYAKIQSTLSTAYMTAKTASNTSA 128
Qy 56 TLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEORATNLEGLSSTAYN 125
Db 129 TLDNLSASTTLQAAIDRAASNKRTFDSANQPLVAYNOLKTLTLOSKTTSLEGLSENKYS 188
Qy 126 QIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITANKNNINTLSTINEQKTNA 185
Db 189 SIKNHLKSLFDAGSAIAARTLDP-TMGIVPEVMSVTKANEDIMTAVSKLTEKTNADKF- 246
Qy 186 NSFIKKVTQNNQESFGYGTFTNAN--VQPSNYSFVAFSADVTP-----VNYKYARRT 234
Db 247 NDFEKKPL--SKEKLVSTNDRAHNQEPANWSFAGYSVDLT*TGSTGNSQNLPNWFAQRK 304
Qy 235 VWNGD--EPPSRILANTNSITDVSWIYSLAGTNTKYQFSNPGSTGYLYFPYKLVKA 291
Db 305 VWTSEGOOTGKTALVSPVSATDVSWIYSLAGEGKTYLTSFEYIGPDTAFLIFPYKLVK 364
Qy 292 ADANNVGLQYKLNNGNVQVEP-----ATST-----S 318
Db 365 ADSSVALQYSLNKTSSKLINEPEAKTPTNADQSENGVATTSTTEGRSSSEVLVADEVA 424
Qy 319 ANNTTANPTTPAYDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSN 378
Db 425 AVNEMNPTPTVSDINIAKVTLSGLTFGENTIEFSVPT-----NKVAPMIGNMYLTSS 479
Qy 379 NADKI 383
Db 480 SQGKI 484

RESULT 6
Q9KH14 PRELIMINARY; PRT; 680 AA.
AC Q9KH14;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE ADHESIN PMGA1.3.
GN PMGA1.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91414.1; -.
SQ SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DFE47 CRC64;

Query Match 38.1%; Score 761; DB 2; Length 680;
Best Local Similarity 39.8%; Pred. No. 5.7e-30;
Matches 164; Conservative 71; Mismatches 129; Indels 48; Gaps 7;

Qy 8 DANPNNGOTQLEAARME-----LTDLINAKAMTSLASLDYAKIEASLSAYSAEYVN 60
Db 63 DTNPGGGQGMNATNQELVNAKKAALSDLIGESKTVELYADYAKIKADLTSAVAKTTIS 122
Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEORATNLEGLS 120
Db 123 DSSTSTLDQVKATATSLQTAINTAASDKKFDQNSQLLMAYKVLKLDTLNKKAEIVMSLN 182
```

121 STAYNQIRNLDVLYNKRASSLIITKTLPLNGGTLDDSGNEITTANKNINNL-----ST 173  
183 QEKYSAILSEINAASTAEITVKTLNPV-----NLPVVAALNAENTKILFAIKEEK 236  
174 INEOKTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNVKYARR 233  
237 INSEKSNADLFANYOLYK-----DRTKLMSEGSNTKPGNISFVAYASDIASPNWFAQR 293  
234 TWNGD-----EPSRILANTNSITDVSWIYSLAGTNTKYQFSFNSGPGTGYLYFPYKL 288  
294 TWTADSTWTSPLPNLONASPLTDSWIYTLSTGAKYTLITDYGPGTGYLYFPYKL 353  
289 VKAADANNVGLQYKLNGNGVQVEF-----ATSTGANTTANPTTAV 330  
354 VKTSD--KVGLQYKLNQADPAIVQSEAAATASAPAEATDGRQESAETATANEKVNPMPSV 411  
331 DEIKVAKIVLSGLRGQNTIELSVPTGCGNMKNKVPMTGNIYLSNENNAK 382  
412 NTINAKVTLSNLKFGSNTIEFSVPMQDQNMKNKVPMTGNIYLSNENNAK 463

RESULT 7  
Q49498 ID Q49498 PRELIMINARY; PRT: 702 AA.  
AC Q49498;  
DT 01-NOV-1996 (TremBLrel. 01, Created)  
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)  
DE 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE PMGAL.3 PROTEIN PRECURSOR.  
GN PMGAL.3.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN [1]  
RC STRAIN=56;  
RX MEDLINE=95010739; PubMed=7925999;  
RA Markham P.F., Giew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
RA Browning G.F., Whitehead K.G., Walker I.D.;  
RT "The organisation of the multigene family which encodes the major cell  
RT surface protein, pmga, of Mycoplasma gallisepticum.";  
RL FEBS Lett. 352:347-352(1994).  
DR EMBL; L28424; AAA62417.1; -  
KW SIGNAL.  
FT SIGNAL.  
SQ SEQUENCE 1 25 POTENTIAL.  
SEQUENCE 702 AA; 75537 MW; 273E8915FE57B9F CRC64;

Query Match 37.7%; Score 754; DB 2; Length 702;  
Best Local Similarity 41.6%; Pred. No. 1.3e-29;  
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

-QY 10 NPNGQT-----QLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVNNLNA 65  
Db 69 NPNGTTPPEQQAARAKTLLDGLTENTVALYADYAKIQSTLSTAYMTAKTASENTSA 128  
QY 66 TLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTATNLEGLSSTAYN 125  
Db 129 TLENRSASTTLQAAIDKANDKRVDSVNOPLVAAYNLTKSKTTSLEGLSENKYG 188  
QY 126 QIRNLDVLYNKRASSLIITKTLPLNGGTLDDSGNEITTANKNINNL-----TINEOKTNA 181  
Db 189 GIKNHLKSLDFTGSAITAKTLTDSGTGRTLEKYN-----ANNIGKMAISPESLKKWKGN 245  
QY 182 DALNSIFIKKVIQNEQSFVGTFTNANVOPSNYSFVAFSADVTP-----VNYKYARVTV 236  
Db 246 DKF-NEPEKNPLSEKELKSDTAHQEQPANWFAIYSDVLTNSQNLPNWFAQRKVW 304  
QY 237 NGD--EPSRILANTN--SITDVSWIYSLAGTNTKYQFSFNSGPGTGYLYFPYKLKRAAD 293  
Db 305 TSENQPGKALVSSPVSATDVSWIYSLAGTNTKYQFSFNSGPGTGYLYFPYKLKRAAD 364

294 ANNVLQYKLNGNGVQVEF-----ATST-----SAN 320  
Db 365 SSSVALQYSLNKTSSKLNFKPAETVSTNTQSENEVATTSTTEARSYKVLVADEAATS 424  
QY 321 NTTANPTPAVDKIVAKIVLSGLRGQNTIELSVPTGCGNMKNKVPMTGNIYLSNENNA 380  
Db 425 NNNHHTPTVSDIINIAKVTLSGLTFGENTIEFSVPEG-----KVAPMIGNMYLTSNSESQ 479  
QY 381 DK1 383  
Db 480 VK1 482

RESULT 8  
Q92IDI ID Q92IDI PRELIMINARY; PRT: 645 AA.  
AC Q92IDI;  
DT 01-MAY-1999 (TremBLrel. 10, Created)  
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TremBLrel. 10, Last annotation update)  
DE M9 PROTEIN.  
GN M9.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN [1]  
RC STRAIN=PG31; ATCC19610;  
RX MEDLINE=99003182; PubMed=9784576;  
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;  
RT "A protein (M9) associated with monoclonal antibody-mediated  
RT agglutination of Mycoplasma gallisepticum is a member of the pmga  
RT family.";  
RL Infect. Immun. 66:5570-5575(1998).  
DR EMBL; AF032890; AAC69269.1; -  
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAB055 CRC64;

Query Match 37.3%; Score 746; DB 2; Length 645;  
Best Local Similarity 42.3%; Pred. No. 2.9e-29;  
Matches 172; Conservative 64; Mismatches 123; Indels 48; Gaps 11;

QY 8 DANPNNGQ-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 60  
Db 44 DTNPGDQGMNNAASOELAAARMGLTIVFDSKAKNLGLYVDYKKTQDTLTKAYDAKTVL 103  
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTATNLEGLS 120  
Db 104 DNSSTTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKVYELKTTLSNETATLAPYA 163  
QY 121 STAYNQIRNLDVLYNKRASSLIITKTLPLNGGTLDDSGNEITTANKNINNL-----STINEOK 178  
Db 164 AAQYAGIKMHLGSLYDAGAKITTKTLEPEVGGDP--LTADVVMMANTKIVEAIKDEVLPQK 222  
QY 179 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NRY 230  
Db 223 ENATKADSFVKVLYKEKITGVEEAHN--RAQPNYSFVGVSDITGTVTQTSIPNWDY 281  
QY 231 ARTVW--NGDEPSSRILANT-----NSITDVSWIYSLAGTNTKYQFSFNSGPGTGY 281  
Db 282 AORTIFTNGDEP--RSISNTPADGQTMVQPLSNYSWIYSLAGTNTKYQFSFNSGPGTGY 339  
QY 282 LYFPYKLKRAADANNVGLQYKLNGNGVQVEFATSTSA-----NNTANPTPAVDEIKVA 336  
Db 340 LYFPYKLKRAADANNVGLQYKLNGNGVQVEFATSTSA-----NNTANPTPAVDEIKVA 391  
QY 337 KIVLSGLRGQNTIELSVPTGCGNMKNKVPMTGNIYLSNENNAK 383  
Db 392 KVTLANLKFGSKIEFSVPA-----EKVSPMIGNMYLTSNSENNAK 433

```
Q92HR9
ID Q92HR9 PRELIMINARY; PRT; 644 AA.
AC Q92HR9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmGA
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL: AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FDE5C CRC64;

Query Match 37.3%; Score 745; DB 2; Length 644;
Best Local Similarity 42.3%; Pred. No. 3.3e-29;
Matches 172; Conservative 65; Mismatches 122; Indels 48; Gaps 11;

Qy 8 DANPNQ-----TQLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
Db 44 DTPGDOGMMNAAQELAAARMGLTTVFDSKAKNLGLYVDYKKTQDTLTAKYDAAKTVL 103
Qy 61 NNLNATLEOLKMAKTNLESAINQANTOKTFDNEHPNLVEAYKALKTTLEORATNLEGLS 120
Db 104 DNSSTQNLNEAKRLETAIRTAATSKQTFDEQHAELVKVYEELKTTLSNETATLAPYA 163
Qy 121 STAYNQIRNLVDLYNKASSLITKTLPLNGTLLDSNEITANKNINNTL--STINBQK 178
Db 164 AAQYAGIKMHLGSLYDAGKAITTKTLEPEVGDPLTASAVMMANTKIVEAIKDEVLPQK 222
Qy 179 TNADALNSFTKKYQIONNEQSFVGTFTNANVPQSNYSFVAFSADVTPV-----NKKY 230
Db 223 ENATKLADSFVKQVLVKEKITGVEEAHKS-QPANYSFVGSVDITGTANGQTSIPNNNY 281
Qy 231 ARRTVW-NGDEPSSRLANT-----NSITDYSWIYSLAGTNTKYQFSNYPSTGY 281
Db 282 AQRFTINGDEP--RSVSNTFPVGDQTMQAQPLSNVSWIYSLAGTCAKYTLEFTYYPSTGY 339
Qy 282 LYFPYKLVKAADANNVQLYKLNNGVQVFEATSTSA-----NNTTANPTPAYDEIKVA 336
Db 340 LYFPYKLVNTSDQVKGLEYKLN-----ATKPSAITGSEQTMNGKTPTVNDINVA 391
Qy 337 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383
Db 392 KVTLANLFGSNKIEFSVP-----VEKVPIMIGNMYLSSSPNNNKKI 433

RESULT 10
Q49495 PRELIMINARY; PRT; 650 AA.
AC Q49495; O08060;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HAEMAGGLUTININ PRECURSOR.
GN PMGAL.1.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.

Q92HR9
ID Q92HR9 PRELIMINARY; PRT; 644 AA.
AC Q92HR9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmGA
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL: AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FDE5C CRC64;

Query Match 37.2%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 3.7e-29;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;

Qy 10 NPNNQ-----TQLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTNNN 62
Db 51 NPGDOGMMNAAQELAAARMGLTTVFDSKAKNLGLYVDYKKTQNTLTAKYDAAKTVLDN 110
Qy 63 LNATLEOLKMAKTNLESAINQANTOKTFDNEHPNLVEAYKALKTTLEORATNLEGLS 122
Db 111 SSSTQNLNEAKRLETAIRTAATSKQTFDEQHAELVKVYEELKTTLSNETATLAPYA 170
Qy 123 AYNQIRNLVDLYNKASSLITKTLPLNGTLLDSNEITANKNINNTL--STINBQKN 180
Db 171 QYAGIKMHLGSLYDAGKAITTKTLEPEVGDPLTASAVMMANTKIVEAIKDEVLPQKN 229
Qy 181 ADALNSFTKKYQIONNEQSFVGTFTNANVPQSNYSFVAFSADVTPV-----NKKYAR 232
Db 230 ATKLADSFVKQVLVKEKITGVEEAHN-KAQPANYSFVGSVDITGTANGQTSIPNNNYAQ 288
Qy 233 RTVW-NGDEPSSRLANT-----NSITDYSWIYSLAGTNTKYQFSNYPSTGYLY 283
Db 289 RTFTINGDEP--RSVSNTFPVGDQTMQAQPLSNVSWIYSLAGTCAKYTLEFTYYPSTGYLY 346
Qy 284 FPKLVKLVKAADANNVQLYKLNNGVQVFEATSTSANNTAN-PTPAVDKIKVAKTVLGS 342
Db 347 FPKLVNTSDQVKGLEYKLNDA---TEPSAITGNEQTMNGKTPTVNDINVAKVTLAN 402
Qy 343 LRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNNKKI 438

RESULT 11
Q49497 PRELIMINARY; PRT; 649 AA.
ID Q49497;
AC Q49497;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PMGAL.2 PROTEIN PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmGA, of Mycoplasma gallisepticum."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmGA, of Mycoplasma gallisepticum.";
```

RL RBBS Lett. 352:347-352(1994).  
DR EMBL: L28424; AAA62416.1; --  
KW SIGNAL.  
FT SIGNAL.  
-SQ SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;  
  
Query Match 37.2%; Score 743; DB 2; Length 649;  
Best Local Similarity 42.3%; Pred. No. 4.1e-29;  
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;  
  
QY 8 DANPNNGQ-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60  
DB 48 DTNPGDGGMMNAASQELAAARMGLTTTFDSKAKNLGLYVDYKKTQNTLTAYDAAKTVL 107  
QY 61 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120  
DB 108 DNSSSTTONLEAKRLETAIRTAATSKOTFDEQHAELVKYKELKTTLSNETATLAPYA 167  
QY 121 STAYNOIRNRLVDLYNKASSLITTKTLDPLNGGTLDSNEITTANKNINNTL--STINEQK 178  
DB 168 DAQYAGIKMHLGLYDAGKAITTKTLEPEVGGP-LTASAVMMANTKIVEAIKDEVLNPQK 226  
QY 179 TNADALSNSFKKVTQNNNEQSFVGTFTNANVQPSNYSFVAFSADVTVP-----NKY 230  
DB 227 ENATKLADSFVKQVLVEKITGVEEHN-KAQPANYSFVGVSDVTGTTGQTSIPNWDY 285  
QY 231 ARRTVW-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 281  
DB 286 AORTIFTSNDEP--RSISNTPADGQTMQAPLSNVSWIYSLAGTNTKYQFSFNSYGPSTGY 343  
QY 282 LVFPYKLVKAADANNVQYKLNNGVQOQVEFATSTSA-----NNTTANPTPAVDKIVA 336  
DB 344 LVFPYKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDQTMNGKTPVTVDINVA 395  
QY 337 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383  
DB 396 KVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437  
  
RESULT 12  
QY9468 PRELIMINARY; PRT; 647 AA.  
AC Q49468; Q53303;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE HEMAGGLUTININ HOMOLOG PRECURSOR.  
-GN PMGAL.2.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
-RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-93162830; PubMed-8432610;  
RA Markham P.F., Glew M.D., Whithear K.G., Walker I.D.;  
RT "Molecular cloning of a member of the gene family that encodes pmGA, a  
RT hemagglutinin of Mycoplasma gallisepticum."  
RL Infect. Immun. 61:903-909(1993).  
DR EMBL: M83178; AAA02996.1; --  
DR EMBL: S55216; AAB2397.2; --  
KW Signal.  
FT CHAIN  
FT CHAIN  
SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;  
  
Query Match 35.7%; Score 713; DB 2; Length 647;  
Best Local Similarity 41.1%; Pred. No. 1.2e-27;  
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;  
  
QY 8 DANPNNGQ-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60

DB 48 DTNPGDGGMMNAASQELAAARMGLTTTFDSKAKNLGLYVDYKKTQNTLTAYDAAKTVL 107  
QY 61 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120  
DB 108 DNSSSTTONLEAKRLETAIRTAATSKOTFDEQHAELVKYKELKTTLSNETATLAPYA 167  
QY 121 STAYNOIRNRLVDLYNKASSLITTKTLDPLNGGTLDSNEITTANKNINNTL--STINEQK 178  
DB 168 DAQYAGIKMHLGLYDAGKAITTKTLEPEVGGP-LTASAVMMANTKIVEAIKDEVLNPQK 226  
QY 179 TNADALSNSFKKVTQNNNEQSFVGTFTNANVQPSNYSFVAFSADVTVP-----NKY 230  
DB 227 ENATKLADSFVKQVLVEKITGVEEHN-KAQPANYSFVGVSDVTGTTGQTSIPNWDY 285  
QY 231 ARRTVW-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 281  
DB 286 AORTIFTSNDEP--RSISNTPADGQTMQAPLSNVSWIYSLAGTNTKYQFSFNSYGPSTGY 343  
QY 282 LVFPYKLVKAADANNVQYKLNNGVQOQVEFATSTSA-----NNTTANPTPAVDKIVA 336  
DB 344 LVFPYKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDQTMNGKTPVTVDINVA 395  
QY 337 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383  
DB 396 KVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437  
  
RESULT 13  
QY9KH15 PRELIMINARY; PRT; 656 AA.  
AC Q9KH15;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE ADHESIN PMGAL.2.  
GN PMGAL.2.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
-RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=HS;  
RA Shen Q.C., Bi D.R., Weng C.J.;  
RT "Sequence analysis of the pmGA multigene family of Mycoplasma  
RT gallisepticum strain HS."  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF275312; AAF91413.1; --  
SQ SEQUENCE 656 AA; 70875 MW; 58EA7E075FC617E1 CRC64;  
  
Query Match 35.2%; Score 704; DB 2; Length 656;  
Best Local Similarity 41.0%; Pred. No. 3.4e-27;  
Matches 166; Conservative 67; Mismatches 124; Indels 48; Gaps 12;  
  
QY 8 DANPNNGQ-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60  
DB 63 DINPGGQNMDSAAQELTAARTALTSLASKANVEMVSDYAKIQNTLIAAYTTAEQTS 122  
QY 61 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120  
DB 123 QNSSATLEQVKNATSAQTANTANSNKRQKQFDQHSNLLMSYKMLMATLAKKETAATVTLK 182  
QY 121 STAYNOIRNRLVDLYNKASSLITTKTLDPLNGGTLDSNEITTANKNINNTL--TINEQK 178  
DB 183 DPKYSAILDQINGVSSGGEELVQHTLPVS-GIVPAANTITEETIKIEEIVSEKTLQDQK 241  
QY 179 TNADALSNSFKKVTQNNNEQSFVGTFTNANVQPSNYSFVAFSADVTVP-----NKY 230  
DB 242 NNAOFANYQSTLTKTLENVEDA-----KMGQAPANYSFVGVSDVTGTSQGETTIP 295  
QY 227 NYKYARTVWNGDEPSSRLANTNS-----ITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 278

Db 296 MNFAQRAIFTSGNPTKVATTTTGGDQSTAKPLSDVSWIYSLAGTAKYTLLEFTYIGPS 355  
Qy 279 TGYLFFPKLVKAADANNVGLQYKLN-NGNQVQVEFATSTANNTTANPTPAVDEIKVAK 337  
Db 356 TGWLYFPKLVKAND--DVGLOKLSNETLPIIFGGT-----TTNGPAATVENINAK 409  
Qy 338 IVLSGLREFGONTIELSVPTGEGNMKNKVPAMIGNIYLSNENNAK 382  
Db 410 VRLTGLAFCKNTIEFVSP-----MSKVAPMIGNMYITSSDTETNK 449  
RESULT 14  
ID O05122 PRELIMINARY; PRT; 703 AA.  
AC O05122;  
DT 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE HAEMAGGLUTININ.  
GN PMGAL.9.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S6;  
RX MEDLINE=95010739; PubMed=7925999;  
RA Markham P.F.; Glew M.D.; Sykes J.E.; Bowden T.R.; Pollocks T.D.;  
Browning G.F.; Whithear K.G.; Walker I.D.;  
RT "The organisation of the multigene family which encodes the major cell  
RT surface protein, pmGA, of Mycoplasma gallisepticum.";  
RL FEBS Lett. 352:347-352(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S6;  
RA Markham P.F.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U90714; AAB50154.1; .  
DR InferPro; IPR002819; HD.  
SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73CB5 CRC64;  
Query Match 34.6%; Score 690.5; DB 2; Length 703;  
Best Local Similarity 35.0%; Pred. No. 1.7e-26;  
Matches 167; Conservative 75; Mismatches 120; Indels 115; Gaps 12;  
Qy 1 CMSTTKKDANP-----NNGQT-----OLEAARMELTDLINAK 32  
Db 26 CTSATIPTLPTNPPEKPDMPNPPSGMNGGNTNPGMDTAQELASAKAALTTLTNR 85  
Qy 33 AMTLASLDYAKIEASLSAYSEATVNNLNATLEQLKMAKTNLESAINQANTDKTTFD 92  
Db 86 SEKVGLYDYAKIKADLTSAVTAKTSDSSTSLVQVKTATSKLQTAIDKAASQKQFE 145  
Qy 93 NEHPNLVEAYKALKTLEQR-ATNLEGLSSTAYNOIRNLVDLYNKASLTKTLDPNG 151  
Db 146 QHKDLLMPYSELKTTLSQKNATVL--LNQPKYSAILNKINSIYAQGEVWIRLTDPSV- 202  
Qy 152 GTLDSNEITTANKNINNTLS--TINEOKTNADALSNSFIKVIQNEQSFVGTFTNANV 209  
Db 203 GAIPTAASITKVNDEINKAISENQLKPKDNADAFANYQFFKL---DKTKINGMSTNMK 259  
Qy 210 QPSNYSFVAFSADYTPV-----NYKYARTVWNGDEPSSRILANTNS----- 251  
Db 260 QPQNTYFVGYSGVTGMSGQTTIPNWNFAORIVWSSGAPRAPLASQTETPOAETPPMSA 319  
Qy 252 -----ITDVSWIYSLA 262  
Db 320 POGVEPAQQGDSPPKQASEQVSEPTFAAEVQAQOADEQPATSGQPLTIDVSWIYSL 379  
Qy 263 GTNTKYQFSFSGYSGTGYLFPKLVKAADANNVGLQYKLNNGNVQVFEATSTSANNT 322

Db 380 GTDVKYTFTFNFGSPMAYLYFPYKLVKSD--SVGLQYKLNNNPNVALNFGSETNAN-- 435  
Qy 323 TANPTPAVDEIKVAKIVLSGLREFGONTIELSVPTGEGNMKNKVPAMIGNIYLSNENN 379  
Db 436 --GPAASVDNINAKVNLNLANLNFNGENTIEFVSP-----MNKVAPMIGNMYITSDVAN 485  
RESULT 15  
ID Q9L8D5 PRELIMINARY; PRT; 419 AA.  
AC Q9L8D5;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE PMGA-LIKE PROTEIN 9.3 (FRAGMENT).  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F;  
RA Pharr G.T.; Branton S.L.; Hanson L.A.; Minion F.C.; Lott B.D.;  
May J.D.; Hughlett M.B.;  
RT "A novel pmGA-like gene from the F-strain (vaccine strain) of  
RT Mycoplasma gallisepticum.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF210770; AAF29525.1; .  
FT NON\_TER 419 419  
SQ SEQUENCE 419 AA; 45072 MW; FE5EBE37F2DB3B0C CRC64;  
Query Match 31.0%; Score 619.5; DB 2; Length 419;  
Best Local Similarity 40.1%; Pred. No. 2.9e-23;  
Matches 149; Conservative 58; Mismatches 122; Indels 43; Gaps 11;  
Qy 8 DANPNNGQT-----OLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 60  
Db 61 DINPGGQNMDSAAQELTAARTALTSLKANKNANTEMYSDYAKIQNTLIAATTAQTS 120  
Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLLEQRAINLEGLS 120  
Db 121 QNSATLEQVKNATSALQTAINTANSKQKQDQDHSNLLMSYKNLMATLAKKETTMTLK 180  
Qy 121 STAYNQIRNLVDLYNKASLTKTLDPNGTLLDSNEITTANKNINNTLS--TINEOK 178  
Db 181 DPKYSAILDQINGVSCGELVQHTLDPVS-GIVPAANTITEEITKIEEIVISEKTLQDQK 239  
Qy 179 TNADALSN--SFI--KKVIQNEQSFVGTFTNANVQPSNYSFVAFSADYTPV----- 236  
Db 240 NNADQFNTQSTFLDKTKLUNVEDA-----KKMGQPNYSFVGYSDVDTGSGQETTP 293  
Qy 227 NYKYARTVWNGDEPSSRILANTNS-----ITDVSWIYSLAGTNTKYQFSFSGY 278  
Db 294 MNFAQRAIFTSGNPTKVATTTTGGDQSTAKPLSDVSWIYSLAGTAKYTLLEFTYIGPS 353  
Qy 279 TGYLFFPKLVKAADANNVGLQYKLN-NGNQVQVEFATSTANNTTANPTPAVDEIKVAK 337  
Db 354 TGWLYFPKLVKAND--DVGLOKLSNETLPIIFGGT-----TTNGPAATVENINAK 407  
Qy 338 IVLSGLREFGONT 349  
Db 408 VRLTGLAFCKNT 419  
Search completed: June 12, 2002, 10:50:40  
Job time: 203 sec

us-09-147-052-2\_copy\_64\_456.rspt

Thu Jun 13 10:00:48 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:48:38 ; Search time 55.68 Seconds  
(without alignments)  
785.975 Million cell updates/sec

Title: US-09-147-052-4\_COPY\_693\_1086

Perfect score: 2004  
Sequence: 1 GCMSTKKDANPNNGTQLE.....SSNENADKIPGRRPGTEL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2004	100.0	456	18 AAW36050	Hybrid Marek's dis
2	2000	99.8	1086	18 AAW36051	Hybrid Marek's dis
3	1910	95.3	615	15 AAR63230	Mycoplasma gallise
4	1856	92.6	610	15 AAR63229	Mycoplasma gallise
5	1612	80.4	368	14 AAR44493	Mycoplasma gallise
6	1612	80.4	368	15 AAR63227	Mycoplasma gallise
7	1604	80.0	368	16 AAR76955	Mycoplasma gallise
8	1142	57.0	235	10 AAR93646	Amino acid sequenc
9	1142	57.0	235	11 AAR05081	MG-1 antigen. AA
10	1142	57.0	261	11 AAR05082	TMG-1 antigen. A
11	1142	57.0	261	16 AAR79911	M.gallisepticum 26

12	1117	55.7	261	10 AAR93959	Amino acid (AA) se
13	806	40.2	661	15 AAR63226	Mycoplasma gallise
14	806	40.2	661	16 AAR79910	M.gallisepticum 66
15	713	35.6	647	15 AAW11978	Mycobacterium gall
16	700.5	35.0	648	15 AAR56973	PMGA 1.2 protein o
17	386.5	19.3	183	10 AAR93649	Amino acid sequenc
18	386.5	19.3	183	11 AAR06439	MG-4 antigen. AA
19	284.5	14.2	219	10 AAR93648	Amino acid sequenc
20	284.5	14.2	219	11 AAR06438	MG-3 antigen. AA
21	176.5	8.8	6281	22 AAU37403	Staphylococcus aur
22	176	8.8	1095	22 AAG83030	S. epidermidis ope
23	171	8.5	2086	22 AAU34143	Staphylococcus aur
24	171	8.5	5795	22 AAU37017	Staphylococcus aur
25	162	8.1	1029	22 AAU34389	Staphylococcus aur
26	162	8.1	1048	22 AAU37490	Staphylococcus aur
27	159	7.9	933	21 AAY58435	Staphylococcus aur
28	159	7.9	933	22 AAB89508	Staphylococcus aur
29	159	7.9	936	18 AAM98801	Staphylococcus aur
30	158	7.9	2434	22 AAU34339	Staphylococcus aur
31	156	7.8	807	21 AAB18311	Plasmodium falcipa
32	154.5	7.7	682	17 AAR95273	Nisin nisp gene pr
33	153.5	7.7	2478	22 AAU34320	Staphylococcus aur
34	153.5	7.7	2478	22 AAU37374	Staphylococcus aur
35	151.5	7.6	1237	18 AAW55640	H. pylori ORF 04ep
36	151.5	7.6	1237	20 AAY17187	H. pylori outer me
37	151	7.5	5024	22 AAG82935	S. epidermidis ope
38	145.5	7.3	2437	22 AAU34338	Staphylococcus aur
39	145	7.2	1215	22 AAU34412	Staphylococcus aur
40	145	7.2	1269	22 AAU37520	Staphylococcus aur
41	143.5	7.2	837	22 AAU34387	Staphylococcus aur
42	143.5	7.2	875	22 AAU37487	Staphylococcus aur
43	143	7.1	3158	22 AAU37018	Staphylococcus aur
44	142	7.1	135	15 AAR56974	Partial PMGA 1.3 p
45	142	7.1	135	15 AAW11979	Mycobacterium gall

#### ALIGNMENTS

RESULT 1

AAW36050  
ID AAW36050 standard; Protein; 456 AA.

AC AAW36050;

DT 15-JUL-1998 (first entry)

DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;  
antigen; vaccine; poultry.

XX Chimeric - Marek's disease gammaherpesvirus.

OS Chimeric - Mycoplasma gallisepticum.

FH Key Location/Qualifiers

FT Region 1..64

FT Region /note= "derived from Marek's disease virus gB protein"

FT Region 65..456

FT Region /note= "derived from M. gallisepticum antigenic protein"

XX WO9736924-A1.

XX PD 09-OCT-1997.

XX PF 28-MAR-1997; 97WO-JP01084.

XX PR 29-MAR-1996; 96JP-0103548.

XX PA (JAPG ) NIPPON ZEON KK.

XX PI Saito S, Tsuzaki Y, Yanagida N;

XX

DR WPI: 1997-503046/46.  
XX N-PSDB; AAT96595.  
XX Fusion protein comprising herpes virus outer membrane protein and  
PT antigenic polypeptide - for prevention of infection by Mycoplasma  
PT gallisepticum, especially in poultry  
XX  
PS Disclosure: Page 16-19; 51pp; Japanese.  
XX  
XX This sequence represents the chimeric protein 40 K-S which comprises a  
CC fragment of the Marek's disease virus outer membrane protein gB fused  
CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
CC protein can be used in recombinant live vaccines for prevention of  
CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
CC protein shows antigenicity in poultry.  
XX  
SQ Sequence 456 AA;  
  
Query Match 100.0%; Score 2004; DB 18; Length 456;  
Best Local Similarity 100.0%; Pred. No. 1.3e-128;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GCMSTTKDANPNNGOTLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEV 60  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
63 gcmsttkdanpnngqtleaarmeltdlinakamtlaslqdyakleaslsayseae 122  
QY 61 NNNLNLATLEQLKMAKTNLESAINQANTKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
123 nnnlnatleqlkmaktnlesainqantkttdnehpnlveaykalkttleqratnle 182  
QY 121 SSTAYNQIRNNLDVLYNKASSLITKTLDPNGTLLDSNEITTANKNINTLSTINEQKT 180  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
183 sstaynqirnnldvlynkasslittkldplnggtllidsneittanknintlstineq 242  
QY 181 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGD 240  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
243 nadalsnfikkvionneqsfvgtftnnavqpsnyfsvafsadvtpvnykyarrtwng 302  
QY 241 EPSSRLANTNSITDVSWIYSLAGTNKYQFSFNYPSTGYLYFPYKLVKAADANNVGL 300  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
303 epsrllantnsitdvswiyslagtnkyqfsfnypstgylfpyklvkaadannvg 362  
QY 301 QYKLNNGNVQVEFATSTANNTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEG 360  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
363 qykingnvqqvefatstannctantpavdeikvakivlsgrlrgntielsvptgeg 422  
QY 361 NMKVAPMIGNIYLSNENNAKIPGRRPGTFL 394  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
423 nmkvapmignylsnennadkipgyrrpgtfl 456  
  
RESULT 2  
ID AAW36051  
XX AAW36051 standard; Protein; 1086 AA.  
AC AAW36051;  
XX  
DT 15-JUL-1998 (first entry)  
XX  
DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.  
XX  
KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;  
KW antigen; vaccine; poultry.  
XX  
OS Chimeric - Marek's disease gammaherpesvirus.  
OS Chimeric - Mycoplasma gallisepticum.  
XX  
FH Key Location/Qualifiers  
FT Region 1...672  
FT /note= "derived from Marek's disease virus gB protein"  
FT 693..1086  
FT /note= "derived from M. gallisepticum antigen"

XX WO9736924-A1.  
PN  
XX  
PD 09-OCT-1997.  
XX  
PF 28-MAR-1997; 97WO-JP01084.  
XX  
PR 29-MAR-1996; 96JP-0103548.  
XX  
XX (JAPG ) NIPPON ZEON KK.  
PA  
XX Saito S, Tsuzaki Y, Yanagida N;  
PI  
XX WPI: 1997-503046/46.  
DR N-PSDB; AAT96596.  
XX  
XX Fusion protein comprising herpes virus outer membrane protein and  
PT antigenic polypeptide - for prevention of infection by Mycoplasma  
PT gallisepticum, especially in poultry  
XX  
XX Disclosure: Page 22-30; 51pp; Japanese.  
PS  
XX This sequence represents the chimeric protein 40 K-C which comprises a  
CC fragment of the Marek's disease virus outer membrane protein gB fused  
CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
CC protein can be used in recombinant live vaccines for prevention of  
CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
CC protein shows antigenicity in poultry.  
XX  
SQ Sequence 1086 AA;  
  
Query Match 99.8%; Score 2000; DB 18; Length 1086;  
Best Local Similarity 99.7%; Pred. No. 7.5e-128;  
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GCMSTTKDANPNNGOTLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEV 60  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
693 gcmsttkdanpnngqtleaarmeltdlinakamtlaslqdyakleaslsayseae 752  
QY 61 NNNLNLATLEQLKMAKTNLESAINQANTKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
753 nnnlnatleqlkmaktnlesainqantkttdnehpnlveaykalkttleqratnle 812  
QY 121 SSTAYNQIRNNLDVLYNKASSLITKTLDPNGTLLDSNEITTANKNINTLSTINEQKT 180  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
813 sstaynqirnnldvlynkasslittkldplnggtllidsneittanknintlstineq 872  
QY 181 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGD 240  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
873 nadalsnfikkvionneqsfvgtftnnavqpsnyfsvafsadvtpvnykyarrtwng 932  
QY 241 EPSSRLANTNSITDVSWIYSLAGTNKYQFSFNYPSTGYLYFPYKLVKAADANNVGL 300  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
933 epsrllantnsitdvswiyslagtnkyqfsfnypstgylfpyklvkaadannvg 992  
QY 301 QYKLNNGNVQVEFATSTANNTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEG 360  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
993 qykingnvqqvefatstannctantpavdeikvakivlsgrlrgntielsvptgeg 1052  
QY 361 NMKVAPMIGNIYLSNENNAKIPGRRPGTFL 394  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1053 nmkvapmignylsnennadkipgyrrpgtfl 1086  
  
RESULT 3  
ID AAW63230  
XX AAW63230 standard; Protein; 615 AA.  
AC AAW63230;  
XX  
DT 23-JUN-1995 (first entry)  
XX

DE	Mycoplasma gallisepticum antigen (UM-67).
XX	
KW	recombinant avipox virus; live vaccine; mycoplasma antigen.
XX	
OS	Mycoplasma gallisepticum.
XX	
Key	Location/Qualifiers
FT	Protein 1..615
FT	/note= "Trp residues correspond to TGA codons"
XX	
PN	WO9423019-A.
XX	
PD	13-OCT-1994.
XX	
PF	31-MAR-1994; 94WO-JP00541.
XX	
PR	31-MAR-1993; 93JP-0074139.
PR	30-SEP-1993; 93JP-0245625.
XX	
PA	(JAPG ) NIPPON ZEON KK.
PA	(SHIO ) SHIONOGI & CO LTD.
XX	
PI	Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
PI	Saeki S, Saitos, Takahashi K;
XX	
DR	WPI; 1994-333181/41.
DR	N-PSDB; AAQ77857.
XX	
PT	Recombinant avipox virus combining DNA encoding a polypeptide -
PT	exhibiting antigenicity of mycoplasma, useful for the production
PT	of a live vaccine
XX	
PS	Claim 4; Page 87-91; 123pp; Japanese.
XX	
CC	A restriction fragment of the insert of M.gallisepticum genomic clone
CC	PUM-67 containing an open reading frame was sequenced (AAQ77857). The
CC	ORF encodes an antigenic polypeptide (AAK83230). A recombinant avipox
CC	virus comprising the coding sequence can be used as a live vaccine to
CC	protect against infection by Mycoplasma gallisepticum.
XX	
SQ	Sequence 615 AA;
Query Match	95.3%; Score 1910; DB 15; Length 615;
Best Local Similarity	98.2%; Pred. No. 4.7e-122;
Matches	376; Conservative 3; Mismatches 4; Indels 0; Gaps
QY	2 CMSTTKKDANPNNGCQGLEAARHELTDLINAKAMTSLASLDYAKTIPASISSAYSEATVN 61
DB	
QY	27 cmsittkdanpnnggtqlgaemeltlllnakartiaslqdyakleasissayseaeatvn 86
DB	
QY	62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLNVEAYKALKTTLEQRATNLEGLS 121
DB	
QY	87 nlnnatleqlkmaaktulesaingandkttfdnehnplveaykalkttleqratnlegia 146
DB	
QY	122 STAYNQIRNNLDVLYNKASSLTITKLDPLNGGTLDSNEITTANKNNINNTSLTINEQKTN 181
DB	
QY	147 staynqirnnldvlynnassliktktdplnggmldtsneittvnrninntlstineqktn 206
DB	
QY	182 ADALNSFIKVKVTONNEQSPVGTFTNANVQPSNYSFVAFSAOVTPVNYKYARRTVWNGDE 241
DB	
QY	207 adaLnsfIkVkiqneeqsvgtftnanvqpsnyfsafsdvtpvnykyarrtvwnge 266
DB	
QY	242 PSSRLTANTNSITDVSNIYSLACTNFKYQPSFSGNYGPTGYLFFPKLVKAADANNVGLQ 301
DB	
QY	267 pssrltantnsitdvsniyslagntkcyqfsfsgnygptgylyfpyklvkaadannvgliq 326
DB	
QY	302 YKLNGNVQVEFATSTSANNTANPTPAYDEIKVAKIVLSGLRFQONTIELSVPTGEGN 361
DB	
QY	327 yklngnvqvefatstsanntanptpaydeikvakiVlsgrlfgntielsvptgegn 386
DB	
QY	362 MNKVAPMIGNIYLSNENNADKI 384
DB	

QY 242 PSSRLANTNSITDVSWIYSLAGTNTKYQFSNYPSTGHLVPPYKLVKAADANNVGLQ 301  
 |||||  
 Db 267 pssrlantnsitdvswiyslsgtntkyqfsnypstgylfpyklvkaadasnvgllq 326  
 |||||  
 QY 302 YKLNNNGVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGSGN 361  
 |||||  
 Db 327 yklngnvgpvefatstsannttanptpavdeikvakivlsglrfgqntielsvptgern 386  
 |||||  
 QY 362 MNKVAPMIGNIYLSNENNAADK 383  
 |||||  
 Db 387 mnkvapmignmyitssnaeank 408  
 |||||  
 RESULT 5  
 ID AAR44493  
 XX AAR44493 standard; Protein; 368 AA.  
 AC AAR44493;  
 XX  
 DT 16-JUN-1994 (first entry)  
 XX  
 DE Mycoplasma gallisepticum 40kd antigen.  
 XX  
 KW Vaccine; mycoplasma infection; poultry; fowl.  
 XX  
 OS Mycoplasma gallisepticum.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 262  
 FT /note= "corresponds to NNN codon in AAO53419"  
 FT Misc-difference 283  
 FT /note= "corresponds to NNN codon in AAO53419"  
 FT  
 XX W09324646-A.  
 PN  
 XX  
 XX 09-DEC-1993.  
 PD  
 XX  
 PF 28-MAY-1993; 93WO-JP00715.  
 XX  
 PR 29-MAY-1992; 92JP-0138819.  
 XX  
 XX (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI & CO LTD.  
 XX  
 XX Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;  
 XX  
 XX WPI; 1993-405837/50.  
 DR N-PSDB; AAO53419.  
 XX  
 XX Mycoplasma gallisepticum antigen and DNA coding for it - useful  
 PT for vaccination of fowl against mycoplasma infections  
 XX  
 XX Claim 2; Page 23-26; 37pp; Japanese.  
 PS  
 CC The sequence coding for the 40kDa antigen was obtained by PCR  
 CC amplification of M.gallisepticum genomic DNA. The antigen reacts  
 CC with Mycoplasma-immune or Mycoplasma-infected serum and can be used  
 CC as a vaccine to protect fowl from M.gallisepticum infection.  
 XX  
 XX Sequence 368 AA;  
 SQ

Query Match 80.4%; Score 1612; DB 14; Length 368;  
 Best Local Similarity 95.0%; Pred. No. 4.8e-102;  
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 CMSTTKDANPNQTOLEARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 61  
 |||||  
 Db 27 cmsittkdnpnngqtdlgarmeltdlinakartlasldyakielaslsayseaeavn 86  
 |||||  
 QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNVLVEAYKALKTTLEQRATNLEGLS 121  
 |||||

Db 87 mnlntleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnlegla 146  
 QY 122 STAYNQIRNNLVDLNKAASSLITKTLDPNGCTLLDSNEITTANKNINNTLSTINQOKTN 181  
 |||||  
 Db 147 staynqirnnlvdllynnasslitktdlpinggmllidsneittvnnntlnstineqkcn 206  
 |||||  
 QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241  
 |||||  
 Db 207 adalnsfikkvionneqsfvgtftnanvqpsnysfafsadvtpvnykyarrtvxngde 266  
 |||||  
 QY 242 PSSRLANTNSITDVSWIYSLAGTNTKYQFSNYPSTGHLVPPYKLVKAADANNVGLQ 301  
 |||||  
 Db 267 pssrlantnsitdvsxiyslagtntkyqfsnypstgylfpyklvkaadannvgllq 326  
 |||||  
 QY 302 YKLNNNGVQOQVEFATSTANNTTANPTPAVDEIKVAK 338  
 |||||  
 Db 327 yklngnvgqvefatstsannttanptgqmrllkllk 363  
 |||||  
 RESULT 6  
 AAR63227  
 ID AAR63227 standard; Protein; 368 AA.  
 XX  
 AC AAR63227;  
 XX  
 DT 23-JUN-1995 (first entry)  
 XX  
 DE Mycoplasma gallisepticum 40kd antigen.  
 XX  
 KW recombinant avipox virus; live vaccine; mycoplasma 40kd antigen;  
 KW TTM-1.  
 XX  
 OS Mycoplasma gallisepticum.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 262  
 FT /note= "corresponds to a NNN codon"  
 FT Misc-difference 283  
 FT /note= "corresponds to a NNN codon"  
 FT  
 XX W09423019-A.  
 PN  
 XX  
 XX 13-OCT-1994.  
 PD  
 XX 31-MAR-1994; 94WO-JP00541.  
 PF  
 XX 31-MAR-1993; 93JP-0074139.  
 PR 30-SEP-1993; 93JP-0245625.  
 XX  
 XX (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI & CO LTD.  
 XX  
 XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;  
 PI Saeki S, Saitos, Takahashi K;  
 XX  
 XX WPI; 1994-333181/41.  
 DR N-PSDB; AAO77854.  
 XX  
 XX Recombinant avipox virus combining DNA encoding a polypeptide -  
 PT exhibiting antigenicity of mycoplasma, useful for the production  
 PT of a live vaccine  
 XX  
 XX Claim 4; Page 71-74; 123pp; Japanese.  
 PS  
 CC The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding  
 CC for the 40kd antigen of Mycoplasma gallisepticum under the control  
 CC of a synthetic promoter. A 1300 bp restriction fragment containing  
 CC the promoter-ORF sequence was excised and was used in the  
 CC construction of plasmid pNZ7929-R2. This in turn was involved in the  
 CC construction of a recombinant avipox virus vector comprising the  
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from  
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FpV  
 CC sequences. The recombinant avipox virus is useful as a live vaccine

CC to protect against infection by Mycoplasma gallisepticum.

XX  
SQ Sequence 368 AA;

Query Match 80.4%; Score 1612; DB 15; Length 368;  
Best Local Similarity 95.0%; Pred. No. 4.8e-102;  
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEATVN 61  
DB 27 cmsitkdkdanpnngqtqlearmeltdlinakartlasldgyakleaslsayseatvn 86  
QY 62 NNLNATLQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121  
DB 87 nnlnatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146  
QY 122 STAYNQIRNNLVLDYNKASSLITKTLDPNGTLLDSNEITTANKNINNTLSTINEQKTN 181  
DB 147 staynqirnnlvldynnasslittkldplnggmllidsneittvnrnintstineqktn 206  
QY 182 ADALNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241  
DB 207 adalnsfkkviqneqsfvgftfnanvqpsnyfsvafsadvtpvnykyarrtvxngde 266  
QY 242 PSSRLANTNSITDVSWIYSLAGTNTKYQFSNTPGPGTGYLYFPYKLVKAADANNVGLQ 301  
DB 267 psrllantnsitdvsxiyslagtnkyqfsnypgstgylyfpyklvkaadannvqlq 326  
QY 302 YKLNNGNVQVEFATSTANNTTANPTPAVDEIKVAK 338  
DB 327 yklnngnvqvefatstannnttanptqqlmrklilk 363

RESULT 7

AAAR76955  
ID AAR76955 standard; Protein: 368 AA.  
XX  
AC AAR76955;  
XX  
DT 12-MAR-1996 (first entry)  
XX  
DE Mycoplasma gallisepticum antigenic protein TTM-1.  
XX  
KW Antigenic protein; vaccine; poultry; diagnosis; TTM-1.  
XX  
OS Mycoplasma gallisepticum.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 262  
FT /note= "any amino acid"  
FT Misc-difference 283  
FT /note= "any amino acid"  
XX  
PN JP07133295-A.  
XX  
PD 23-MAY-1995.  
XX  
PF 27-AUG-1993; 93JP-0213102.  
XX  
PR 27-AUG-1993; 93JP-0213102.  
XX  
PA (SHIO ) SHIONOGI & CO LTD.  
XX  
XX WPI; 1995-220782/29.  
DR N-PSDB; AAR76955.  
XX  
PT A new antigenic protein which reacts with Mycoplasma gallisepticum -  
PT is useful in a component vaccine for use against poultry infected  
XX with M. gallisepticum.  
XX  
PS Claim 6; Figs 5-6; 33pp; Japanese.  
XX

CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein  
CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum  
CC infectious diseases in poultry, and as a diagnostic agent for  
CC M. gallisepticum.

XX  
SQ Sequence 368 AA;

Query Match 80.0%; Score 1604; DB 16; Length 368;  
Best Local Similarity 94.4%; Pred. No. 1.7e-101;  
Matches 318; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEATVN 61  
DB 27 cmsitkdkdanpnngqtqlearmeltdlinakartlasldgyakleaslsayseatvn 86  
QY 62 NNLNATLQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121  
DB 87 nnlnatleqlkmaktnlesainqantdkttfdnehpnlvqpykalkttleqratnleqla 146  
QY 122 STAYNQIRNNLVLDYNKASSLITKTLDPNGTLLDSNEITTANKNINNTLSTINEQKTN 181  
DB 147 staynqirnnlvldynnasslittkldplnggmllidsneittvnrnintstineqktn 206  
QY 182 ADALNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241  
DB 207 adalnsfkkviqneqsfvgftfnanvqpsnyfsvafsadvtpvnykyarrtvxngde 266  
QY 242 PSSRLANTNSITDVSWIYSLAGTNTKYQFSNTPGPGTGYLYFPYKLVKAADANNVGLQ 301  
DB 267 psrllantnsitdvsxiyslagtnkyqfsnypgstgylyfpyklvkaadannvqlq 326  
QY 302 YKLNNGNVQVEFATSTANNTTANPTPAVDEIKVAK 338  
DB 327 yklnngnvqvefatstannnttanptqqlmrklilk 363

RESULT 8

AAAP93646  
ID AAR93646 standard; protein: 235 AA.  
XX  
AC AAP93646;  
XX  
DT 11-MAY-1990 (first entry)  
XX  
DE Amino acid sequence of Mycoplasma gallisepticum (MGI) polypeptide.  
XX  
KW Mycoplasma gallisepticum; Poultry vaccine; ss;  
XX  
OS Mycoplasma gallisepticum.  
XX  
PN EP345021-A.  
XX  
PD 06-DEC-1989.  
XX  
PF 31-MAY-1989; 89EP-0005441.  
XX  
PR 02-JUN-1988; 88JP-0136343.  
XX  
PA (JAPG ) NIPPON ZEON KK.  
XX (SHIO ) SHIONOGI SEIVAKU KABUSHIKI KAISHA.  
XX  
PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
XX WPI; 1989-358393/49.  
DR N-PSDB; AAN92568.  
XX  
PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
PT vaccines.  
XX  
PS Disclosure; Fig.1a; 3lpp; English.  
XX  
CC This amino acid sequence of MGI is encoded by M1 DNA and elicits an

CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a  
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro  
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.

XX  
 SQ Sequence 235 AA;

Query Match 57.0%; Score 1142; DB 10; Length 235;  
 Best Local Similarity 97.0%; Pred. No. 2.8e-70;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 61  
 |||||  
 Db 1 cmsitkdkanpnngqqlqaarmeltdlinakartilasldyakielasissayseaevt 60  
 |||||  
 QY 62 NNINATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 121  
 |||||  
 Db 61 nninatleolkmaaktulesainganctktdfnehpnlveaykalkttleqratnlegia 120  
 |||||  
 QY 122 STAYNQIRNNLDLYNKASSLITKTLDPNGGTLTLDSEITTTANKNNINNTLSTINQKTN 181  
 |||||  
 Db 121 staynqirnnldlynnasslittktdplnggmltdsneittvnrnintlstineqktn 180  
 |||||  
 QY 182 ADALNSFTKKVIQNNESQSVGTFNNVQPSNYSFVAFSADVTPVNYKYARETV 236  
 |||||  
 Db 181 adalsnftkkvigneqsfvgfttnanvqpsnysvafsadvtpvnykyarrtv 235

#### RESULT 9

AAR05081  
 ID AAR05081 standard; protein; 235 AA.

XX  
 AC AAR05081;

XX  
 DT 08-OCT-1990 (first entry)

XX  
 DE MG-1 antigen.

XX  
 KW Mycoplasma gallisepticum; poultry; vaccine.

XX  
 PN JP02111795-A.

XX  
 PD 24-APR-1990.

XX  
 PF 02-JUN-1989; 89JP-0136343.

XX  
 PR 02-JUN-1989; 89JP-0136343.

XX  
 PA (JAPG ) NIPPON ZEON KK.

XX  
 PA (SHIO ) SHIONOGI KK.

XX  
 DR WPI; 1990-169109/22.

DR  
 N-PSDB; AAQ04686.

XX  
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilises

PT antigen protein of the disease and recombinant vector

PT incorporated with its coding gene.

XX  
 PS Claim 2; Fig 1a; 20pp; Japanese.

XX  
 CC DNA encoding the protein can be inserted into an expression vector  
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 CC be ligated to other DNA to produce fusion proteins with an N-terminal  
 CC bacterial enzyme sequence.

CC See also AAR05081-2 and AAR06437-41.

XX  
 SQ Sequence 235 AA;

Query Match 57.0%; Score 1142; DB 11; Length 235;  
 Best Local Similarity 97.0%; Pred. No. 2.8e-70;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 61  
 |||||  
 Db 1 cmsitkdkanpnngqqlqaarmeltdlinakartilasldyakielasissayseaevt 60  
 |||||  
 QY 62 NNINATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 121  
 |||||  
 Db 61 nninatleolkmaaktulesainganctktdfnehpnlveaykalkttleqratnlegia 120  
 |||||  
 QY 122 STAYNQIRNNLDLYNKASSLITKTLDPNGGTLTLDSEITTTANKNNINNTLSTINQKTN 181  
 |||||  
 Db 121 staynqirnnldlynnasslittktdplnggmltdsneittvnrnintlstineqktn 180  
 |||||  
 QY 182 ADALNSFTKKVIQNNESQSVGTFNNVQPSNYSFVAFSADVTPVNYKYARETV 236  
 |||||  
 Db 181 adalsnftkkvigneqsfvgfttnanvqpsnysvafsadvtpvnykyarrtv 235

#### RESULT 10

AAR05082

ID AAR05082 standard; protein; 261 AA.

XX  
 AC AAR05082;

XX  
 DT 08-OCT-1990 (first entry)

XX  
 DE TMG-1 antigen.

XX  
 KW Mycoplasma gallisepticum; poultry; vaccine.

XX  
 PN JP02111795-A.

XX  
 PD 24-APR-1990.

XX  
 PF 02-JUN-1989; 89JP-0136343.

XX  
 PR 02-JUN-1989; 89JP-0136343.

XX  
 PA (JAPG ) NIPPON ZEON KK.

PA  
 (SHIO ) SHIONOGI KK.

XX  
 DR WPI; 1990-169109/22.

DR  
 N-PSDB; AAQ04687.

XX  
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilises

PT antigen protein of the disease and recombinant vector

PT incorporated with its coding gene.

XX  
 PS Claim 2; Fig 2; 20pp; Japanese.

XX  
 CC DNA encoding the protein can be inserted into an expression vector  
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 CC be ligated to other DNA to produce fusion proteins with an N-terminal  
 CC bacterial enzyme sequence.

CC See also AAR05081 and AAR06437-41.

XX  
 SQ Sequence 261 AA;

Query Match 57.0%; Score 1142; DB 11; Length 261;  
 Best Local Similarity 97.0%; Pred. No. 3.2e-70;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 61  
 |||||  
 Db 27 cmsitkdkanpnngqqlqaarmeltdlinakartilasldyakielasissayseaevt 86  
 |||||

QY 62 NNINATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 121  
 |||||  
 Db 87 nninatleolkmaaktulesainganctktdfnehpnlveaykalkttleqratnlegia 146  
 |||||  
 QY 122 STAYNQIRNNLDLYNKASSLITKTLDPNGGTLTLDSEITTTANKNNINNTLSTINQKTN 181

Db 147 staynqirnlvdlynnasslittkldplngmildsneittvnrnntlstineqktn 206  
 QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236  
 Db 207 adalsnfsfikkvionneqsfvgtftnqpsnysfvaafsdvtpvnykyarrtv 261

RESULT 11  
 AAR79911  
 ID AAR79911 standard; Protein; 261 AA.  
 AC AAR79911;  
 XX 19-JUL-1996 (first entry)  
 DE M.gallisepticum 261 amino acid protein.  
 XX Detection; probe; primer; PCR; amplification; secretion; lung;  
 KW avian chronic respiratory disease; respiratory tract; nasal cavity.  
 XX Mycoplasma gallisepticum.  
 OS  
 XX JP07236498-A.  
 PN 12-SEP-1995.  
 XX 25-FEB-1994; 94JP-0052764.  
 PF 25-FEB-1994; 94JP-0052764.  
 XX (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI & CO LTD.  
 XX WPI; 1995-347462/45.  
 DR N-PSDB; AAT04076.

XX Detection of Mycoplasma gallisepticum - for the quick detection,  
 PT i.e. within one day, of avian chronic respiratory diseases  
 XX Claim 3; Page 10-11; 1lpp; Japanese.  
 XX This is the amino acid sequence of a 261 amino acid protein encoded  
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding  
 CC sequence and the sequence of AAT04075 (encoding a 661 amino acid  
 CC protein) can be used to detect M.gallisepticum using probes based on  
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the  
 CC complement of bases 893-919, 1908-1934 and the complement of bases  
 CC 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA  
 CC encoding this protein. The method using these sequences is faster i.e. is  
 CC able to detect M.gallisepticum, which causes avian chronic respiratory  
 CC diseases, within one day, from avian secretions, washings from the lung,  
 CC respiratory tract, nasal cavity, etc.  
 XX Sequence 261 AA;

Query Match 57.08; Score 1142; DB 16; Length 261;  
 Best Local Similarity 97.08; Pred. No. 3.2e-70;  
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 CMSITTKDANPNNGQTQLEAARMELTDLINAKAMTFLASLDYAKIEASLSAYSEAEVTN 61  
 Db 27 cmsitkddanpnngqtlqaarmeltdlinakarflaslqdyakieslsayseaevt 86  
 QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121  
 Db 87 nlnnatleqlkmaKNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
 QY 122 STAYNQIRNLVDLYNNASSLITKLDPLNGMILDSEITVNRNNTLSTINEQKTN 181  
 Db 147 staynqirnlvdlynnasslittkldplngmildsneittvnrnntlstineqktn 206

QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236  
 Db 207 adalsnfsfikkvionneqsfvgtftnqpsnysfvaafsdvtpvnykyarrtv 261  
 RESULT 12  
 AAP93959  
 ID AAP93959 standard; protein; 261 AA.  
 AC AAP93959;  
 XX 11-MAY-1990 (first entry)  
 DE Amino acid (AA) sequence of TMG-1 polypeptide.  
 XX Mycoplasma gallisepticum; Poultry vaccine; ss;  
 OS Mycoplasma gallisepticum.  
 PN EP345021-A.  
 XX 06-DEC-1989.  
 XX 31-MAY-1989; 89EP-0005441.  
 XX 02-JUN-1988; 88JP-0136343.  
 XX (JAPG ) NIPPON ZEON KK; (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.  
 XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 DR WPI; 1989-358393/49.  
 DR N-PSDB; AAN92574.  
 XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
 PT vaccines.  
 XX Disclosure; Fig.2; 3lpp; English.

XX This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the  
 CC same sequence as that of a polypeptide expressed in Mycoplasma  
 CC gallisepticum in nature. When the corresponding DNA sequence is inserted  
 CC into a recombinant vector used to transform a host the antigen protein  
 CC produced can be used as a vaccine to prevent and diagnose MG infection.  
 XX Sequence 261 AA;

Query Match 55.78; Score 1117; DB 10; Length 261;  
 Best Local Similarity 96.28; Pred. No. 1.6e-68;  
 Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;  
 QY 2 CMSITTKDANPNNGQTQLEAARMELTDLINAKAMTFLASLDYAKIEASLSAYSEAEVTN 61  
 Db 27 cmsitkddanpnngqtlqaarmeltdlinakarflaslqdyakieslsayseaevt 86  
 QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121  
 Db 87 nlnnatleqlkmaKNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
 QY 122 STAYNQIRNLVDLYNNASSLITKLDPLNGMILDSEITVNRNNTLSTINEQKTN 180  
 Db 147 staynqirnlvdlynnasslittkldplngmildsneittvnrn-ntlstineqkt 205  
 QY 181 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236  
 Db 206 nadalsnfsfikkvionneqsfvgtftnqpsnysfvaafsdvtpvnykyarrtv 261  
 RESULT 13  
 AAR63226  
 ID AAR63226 standard; Protein; 661 AA.  
 XX

20

20

RESULT	15	
AAW11978		
ID	AAW11978 standard; Protein; 647 AA.	
XX		
XX	AAW11978;	
XX		
DT	23-APR-1997 (first entry)	
XX		
DE	Mycobacterium gallisepticum pmGA1.2 adhesin.	
XX		
KW	Adhesin; pmGA; mycoplasma; diagnosis; vaccine; vector;	
XX	respiratory disease; poultry; haemagglutinin.	
XX		
OS	Mycoplasma gallisepticum strain S6.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..25
FT		/label= Sig_peptide
FT		/note= "the signal peptide shows homology to
FT		the pmGA1.3 signal peptide"
XX		
XX	CA2135330-A.	
PN		
XX		
PD	11-MAY-1995.	
XX		
XX	08-NOV-1994; 94CA-2135330.	
XX		
PR	20-APR-1994; 94US-0230312.	
PR	10-NOV-1993; 93AU-0050593.	
XX		
PA	(BROW/) BROWNING G F.	
XX		
PI	Browning GF, Glew MD, Markham PF, Walker ID, Whithear KG;	
XX		
DR	WPI; 1995-241027/32.	
DR	N-PSDB; AAT51531.	
XX		
PT	New promoter region from a Mycoplasma gallisepticum adhesin gene -	
PT	useful when coupled to foreign antigen gene, for prodn. of	

Search completed: June 12, 2002, 10:48:39

Job time: 387 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:49:08 ; Search time 22.56 seconds  
(without alignments)  
426.582 Million cell updates/sec

Title: US-09-147-052-4\_COPY\_693\_1086

Perfect score: 2004

Sequence: 1 GCMSTTKKDANPNNGTQLE.....SSNENADKIPGRRPGTFL 394

-Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

\*Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	95.3	615	2	US-08-525-742-10
2	1856	92.6	610	2	US-08-525-742-8
3	1638	81.7	368	2	US-08-525-742-4
4	1612	80.4	368	1	US-08-185-851A-4
5	806	40.2	661	2	US-08-525-742-2
6	159	7.9	933	3	US-08-293-728-2
7	159	7.9	933	3	US-08-421-868-2
8	136	6.8	1098	4	US-08-923-992A-8
9	135.5	6.8	1002	4	US-09-268-347-24
10	135.5	6.8	1073	4	US-09-541-782-6
11	134	6.7	1004	4	US-09-268-347-30
12	131.5	6.6	1164	4	US-08-923-992A-2
13	131	6.5	1104	4	US-08-923-992A-4
14	129	6.4	2048	4	US-09-268-347-48
15	126.5	6.3	1128	4	US-08-923-992A-6
16	126.5	6.3	1164	4	US-08-923-992A-10
17	125.5	6.3	518	3	US-09-043-123-2
18	125.5	6.3	3788	4	US-09-336-447A-76
19	124.5	6.2	2411	4	US-09-268-347-36
20	124.5	6.1	1536	1	US-09-038-682-2
21	122.5	6.1	1536	1	US-08-302-832-2
22	122.5	6.1	1536	2	US-08-530-198-2
23	122.5	6.1	1536	2	US-08-469-880-2
24	122.5	6.1	1536	2	US-08-728-470-2
25	122.5	6.1	1536	2	US-08-617-697-2
26	122.5	6.1	1536	4	US-08-719-641-2
27	122.5	6.1	1565	6	5352450-2

28	120.5	6.0	746	5	PCT-US95-10509-2	Sequence 2, Appli
29	120.5	6.0	1104	4	US-09-268-347-28	Sequence 28, Appl
30	120.5	6.0	1104	4	US-09-268-347-34	Sequence 34, Appl
31	119.5	6.0	984	1	US-08-242-932-2	Sequence 2, Appli
32	119.5	6.0	984	1	US-08-714-481-2	Sequence 2, Appli
33	119.5	6.0	984	5	PCT-US95-06111-2	Sequence 2, Appli
34	119.5	6.0	2314	4	US-09-268-347-49	Sequence 49, Appl
35	119	5.9	2354	4	US-09-268-347-47	Sequence 47, Appl
36	118.5	5.9	865	2	US-08-483-101-4	Sequence 4, Appli
37	118	5.9	1612	1	US-08-189-927-2	Sequence 2, Appli
38	117.5	5.9	889	4	US-09-336-447A-15	Sequence 15, Appl
39	117.5	5.9	1912	1	US-08-409-995-4	Sequence 4, Appli
40	117.5	5.9	1912	3	US-08-685-467-4	Sequence 4, Appli
41	117	5.8	2353	4	US-09-377-155-33	Sequence 33, Appl
42	117	5.8	2353	4	US-08-913-942-4	Sequence 4, Appli
43	117	5.8	2353	4	US-09-669-974-33	Sequence 33, Appl
44	116	5.8	941	4	US-09-336-447A-9	Sequence 9, Appli
45	116	5.8	1338	2	US-08-728-470-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-08-525-742-10  
; Sequence 10, Application US/08525742  
; Patent No. 5871742  
; GENERAL INFORMATION:  
; APPLICANT: Saito, Shuji  
; APPLICANT: Ohkawa, Setsuko  
; APPLICANT: Saeki, Sakiko  
; APPLICANT: Ohsawa, Ikuroh  
; APPLICANT: Funato, Hirono  
; APPLICANT: Iritani, Yoshikazu  
; APPLICANT: Aoyama, Shigeml  
; APPLICANT: Takahashi, Kiyoohito  
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
; TITLE OF INVENTION: AS USE THEREOF  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
; STREET: 1725 K Street, Suite 1000  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/525,742  
; FILING DATE: 25-SEP-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 05-074139  
; FILING DATE: 31-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 05-245625  
; FILING DATE: 30-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JF94/00541  
; FILING DATE: 31-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mclelland, Le-Nhung  
; REGISTRATION NUMBER: 31,541  
; REFERENCE/DOCKET NUMBER: 950811  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-659-2930

TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 615 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-10

Query Match 95.3%; Score 1910; DB 2; Length 615;  
Best Local Similarity 98.2%; Pred. No. 3.9e-132; Indels 0; Gaps 0;  
Matches 376; Conservative 3; Mismatches 4;

QY 2 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTDLASLDYAKIEASLSAYSAETVN 61  
DB 27 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTDLASLDYAKIEASLSAYSAETVN 86  
QY 62 NNLNATLEQLKMAKTNLESAIQAANTDKTTDFNEHPNLVEAYKALKTTLEQRTNLEGLS 121  
DB 87 NNLNATLEQLKMAKTNLESAIQAANTDKTTDFNEHPNLVEAYKALKTTLEQRTNLEGLA 146  
QY 122 STAYNOIRNLDVLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 181  
DB 147 STAYNOIRNLDVLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 206  
QY 182 ADALNSFTKVKVIONNEQSFVGTFTNANVQPSNYSEVAFSADVTPVNYKYARTVWNGDE 241  
DB 207 ADALNSFTKVKVIONNEQSFVGTFTNANVQPSNYSEVAFSADVTPVNYKYARTVWNGDE 266  
QY 242 PSSRILANTNSITDVSWMYSLAGTNTKYQSFNSYGPSTGYLFFPKYKLYKAADANNVGLQ 301  
DB 267 PSSRILANTNSITDVSWMYSLAGTNTKYQSFNSYGPSTGYLFFPKYKLYKAADANNVGLQ 326  
QY 302 YKLNGNVQVVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 361  
DB 327 YKLNGNVQVVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386  
QY 362 MNKVAPMIGNIYLSNENNAKDI 384  
DB 387 MNKVAPMIGNIYLSNENNAKDI 409

RESULT 2  
US-08-525-742-8  
Sequence 8, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Ohsawa, Ikuroh  
APPLICANT: Funato, Hirono  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESSEE: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
APPLICATION DATA: JP 05-245625  
FILING DATE: 30-SEP-1993  
APPLICATION DATA: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mclelland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-8

Query Match 92.6%; Score 1856; DB 2; Length 610;  
Best Local Similarity 95.5%; Pred. No. 3.5e-128; Indels 0; Gaps 0;  
Matches 365; Conservative 10; Mismatches 7;

QY 2 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTDLASLDYAKIEASLSAYSAETVN 61  
DB 27 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTDLASLDYAKIEASLSAYSAETVN 86  
QY 62 NNLNATLEQLKMAKTNLESAIQAANTDKTTDFNEHPNLVEAYKALKTTLEQRTNLEGLS 121  
DB 87 NNLNATLEQLKMAKTNLESAIQAANTDKTTDFNEHPNLVEAYKALKTTLEQRTNLEGLS 146  
QY 122 STAYNOIRNLDVLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 181  
DB 147 STAYNOIRNLDVLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 206  
QY 182 ADALNSFTKVKVIONNEQSFVGTFTNANVQPSNYSEVAFSADVTPVNYKYARTVWNGDE 241  
DB 207 ADALNSFTKVKVIONNEQSFVGTFTNANVQPSNYSEVAFSADVTPVNYKYARTVWNGDE 266  
QY 242 PSSRILANTNSITDVSWMYSLAGTNTKYQSFNSYGPSTGYLFFPKYKLYKAADANNVGLQ 301  
DB 267 PSSRILANTNSITDVSWMYSLAGTNTKYQSFNSYGPSTGYLFFPKYKLYKAADANNVGLQ 326  
QY 302 YKLNGNVQVVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 361  
DB 327 YKLNGNVQVVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386  
QY 362 MNKVAPMIGNIYLSNENNAKDI 383  
DB 387 MNKVAPMIGNIYLSNENNAKDI 408

RESULT 3  
US-08-525-742-4  
Sequence 4, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Ohsawa, Ikuroh  
APPLICANT: Funato, Hirono

APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi  
APPLICANT: Takahashi, Kiyochito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESSEE: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/525,742  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: McLeLland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 368 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-4

Query Match 81.7%; Score 1638; DB 2; Length 368;  
Best Local Similarity 95.5%; Pred. No. 1.6e-112; Indels 0; Gaps 0;  
Matches 322; Conservative -6; Mismatches 9;

Qy 2 CMSITTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61  
Db 27 CMSITTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86  
Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121  
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
Qy 122 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLDSNEITTTANKNINNTLSTINEQKTN 181  
Db 147 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLDSNEITTTANKNINNTLSTINEQKTN 206  
Qy 182 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTWNGDE 241  
Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTWNGDE 266  
Qy 242 PSSRLANTNSTIDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 301  
Db 267 PSSRLANTNSTIDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 326

Qy 302 YKLNGNVQVEFATSTSANNTTANPTPAVDEIKVAK 338  
Db 327 YKLNGNVQVEFATSTSANNTTANPTQOLMLRLKLLK 363

## RESULT 4

US-08-185-851A-4  
Sequence 4, Application US/08185851A  
Patent No. 5489430  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Fujisawa, Ayumi  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi  
TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene  
TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As  
TITLE OF INVENTION: Well As Vaccines Utilizing the Same  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &  
ADDRESSEE: Naughton  
STREET: 1725 K Street, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0  
SOFTWARE: ASCII from Word Perfect version 5.1  
CURRENT APPLICATION DATA: US/08/185,851A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Theresa M. Stevens-Smith  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: PO-8-A930918  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-887-0357  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 368 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-185-851A-4

Query Match 80.4%; Score 1612; DB 1; Length 368;  
Best Local Similarity 95.0%; Pred. No. 1.3e-110;  
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CMSITTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61  
Db 27 CMSITTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86  
Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121  
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
Qy 122 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLDSNEITTTANKNINNTLSTINEQKTN 181  
Db 147 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLDSNEITTTANKNINNTLSTINEQKTN 206  
Qy 182 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTWNGDE 241  
Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTWNGDE 266



Qy 208 ANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDEPSSRILANTNSITDVSIIYSLAG-TN 266  
Db 292 AKVPB-----IMAGDQ-----VLANGVIDSGNVIYFTDYN 324  
Qy 267 TKYQFSNYPSPGTYLFPYKLVKAADANNVQYKLNNGNVQVQVEFATSTANNTTAN 326  
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362  
Qy 327 PTPAVDEIKVAKI-----VLSGLRFQNTIELSVPTGEGNMKNKVPAMI-GNIYLSNE 378  
Db 363 KTVLVDEYKGYKFNLSIKGTIDQIDKTNNTYRQTIYVPSGDNVIAPLVGLNKPNTDS 422  
Qy 379 N 379  
Db 423 N 423

RESULT 7  
US-09-421-868-2  
; Sequence 2, Application US/09421868  
; Patent No. 6177084  
; GENERAL INFORMATION:  
; APPLICANT: Foster, Timothy J.  
; APPLICANT: Mcdevitt, Damien L.  
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene  
; FILE REFERENCE: 05344.105011  
; CURRENT APPLICATION NUMBER: US/09/421.868  
; CURRENT FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 08/293,728  
; PRIOR FILING DATE: 1994-08-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 933  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-421-868-2

Query Match 7.9%; Score 159; DB 4; Length 933;  
Best Local Similarity 22.8%; Pred. No. 0.0011;  
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

Qy 3 MSITKKDANPNNGTQL--EAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATV 60  
Db 73 VSDTKTSNTNGETSVAQNPAQQTQSSSTNAT-----EETPVTGEATTT 120  
Qy 61 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLNVEAYKALKTTLEQRATNLEGL 120  
Db 121 TTNQANTPATTOSSNTNAEELVNO-TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172  
Qy 121 SST-----AYNQIRNNLVLYNK-----ASSLTKTLDPLNGTLL 156  
Db 173 STTQDTSFTEATPSNNESAQSDTASNDKDVNVQAVNTSAPRMFASLAADAAPAAAGTDI 232  
Qy 157 DSNETTTANKNNLTSTINEQ----KTN-ADALNSFIK----KVIONNEQSFGVGTFTN 207  
Db 233 -TNQLTNVTVIDSGTIVYHQAGVKLVNGYFSPNSAVKGDTKITVPKELNLNGVTST 291  
Qy 208 ANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDEPSSRILANTNSITDVSIIYSLAG-TN 266  
Db 292 AKVPB-----IMAGDQ-----VLANGVIDSGNVIYFTDYN 324  
Qy 267 TKYQFSNYPSPGTYLFPYKLVKAADANNVQYKLNNGNVQVQVEFATSTANNTTAN 326  
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362  
Qy 327 PTPAVDEIKVAKI-----VLSGLRFQNTIELSVPTGEGNMKNKVPAMI-GNIYLSNE 378  
Db 363 KTVLVDEYKGYKFNLSIKGTIDQIDKTNNTYRQTIYVPSGDNVIAPLVGLNKPNTDS 422  
Qy 379 N 379

Db 423 N 423  
RESULT 8  
US-08-923-992A-8  
; Sequence 8, Application US/08923992A  
; Patent No. 6280738  
; GENERAL INFORMATION:  
; APPLICANT: Tai, Joseph Y.  
; APPLICANT: Blake, Milan S.  
; TITLE OF INVENTION: NO. 6280738-IgA Fc Binding Forms of the Group B  
; TITLE OF INVENTION: Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923.992A  
; FILING DATE: 05-SEP-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/024.707  
; FILING DATE: 06-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32.893  
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1098 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-923-992A-8

Query Match 6.8%; Score 136; DB 4; Length 1098;  
Best Local Similarity 22.7%; Pred. No. 0.064;  
Matches 84; Conservative 51; Mismatches 159; Indels 76; Gaps 14;

Qy 5 ITKDDANPNNGTQLEAARMELTDLINA-KAMTLASLDYAKIEASLSAYSE-----AE 58  
Db 192 IRKQAOQPKKEDAEVKVREELGKLFSTKAGLDQIEQHVKKTSSEENTOKVDEHYAN 251  
Qy 59 TVNNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLN-----VEAYKALK 107  
Db 252 SLQNLAKSLEELD KATTN-----EQATQVKNQFLENQAOKLKEIOLPIKETNVKLYKAMS 306  
Qy 108 TTLEQRATNLEGLSSTAYNQIRNNLVNKNASSLITTTLDPLNGTLLDSNEITANKN 167  
Db 307 ESLEQVEKELKHNSA-----NLEDLVAKSKEIVREVEGKLNOSKNLP--ELQQLBEE 357  
Qy 168 INNTLSTINEQ-----KTNADALNSFIKKVI---ONNEQSFVGTFTNANVQPSNYFVA 219  
Db 358 AHSKLLQVVEDFRKKFKTSEQVTPKKRLKRLDLAANNQOKI-----ELTVSPENI---- 408  
Qy 220 PSADVTVPVNYKYARTVWNGDEPSSRILANTNSITDVSIIYSLAGTNTKYQFSFS----- 274  
Db 409 -----TYVEGEDVKFTVAKSDSKTTLDFSDLL-----TKYNPVSVDRIST 449  
Qy 275 NYGPSTG---YLYPPYKLVKAADANNVGLQYKLNNGNVQVQVEFATSTANNTTANP-TPA 330

Thu Jun 13 10:00:50 2002

Db 450 NYKTTNDNHKIAEITIKNLKNSQTVTLKAKDDSGNVVEKFTTITVQKKEKQVPTPE 509  
QY 331 VDEIKVAKIV 340  
Db 510 QKDSKTEKV 519

## RESULT 9

US-09-268-347-24  
; Sequence 24, Application US/09268347  
; Patent No. 6335182  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena M.  
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
; FILE REFERENCE: 1038-860  
; CURRENT APPLICATION NUMBER: US/09/268,347  
; CURRENT FILING DATE: 1999-03-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 1002  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-09-268-347-24

Query Match 6.8%; Score 135.5; DB 4; Length 1002;  
Best Local Similarity 19.2%; Pred. No. 0.061;  
Matches 94; Conservative 60; Mismatches 161; Indels 175; Gaps 21;

QY 1 GCMSTTKDANPNNGOTQLEAARM-----ELTDLINAKAMTLASLDYAKIEASLSAYSE 56  
Db 349 GKLAQKVKLVSAANGIPYKISNVADGTEQDAVSEK--QLKALQD-KQVTLASNAYAN 405  
QY 57 -----AETVNNLNATLEQLMAKTNLESA-----INQANTDKTTFDN 94  
Db 406 GGSADGAGKATQTLGDLNFKFKSTDSSELLNIKAAGDTVTFTPKGVSQVGDGKATIQD 465  
QY 95 -----EHPNVLVEAYKAL-----KTTL----- 110  
Db 466 GAKTTCGLVEASELVDLSNKLGHKVGKDGTDGTDGTHDTLVKSGDKVTLKAGDNLKV 525  
QY 111 EQRATN-----LEGSLSTAYNQIRNNLVLYNKASSLIITK-----TLDP---LNGGTL 156  
Db 526 KQECTNFTYVLRDELTKGVKSEFKDTENGA-----NGASTKITKDGTLTITPANDANGAAAT 581  
QY 157 DSNEITANKNNITLSTINEQKTNADALSNSFIKKVIONNEOSFVGTETFNANVQPSNYS 216  
Db 582 DADKIKVASDGI-----SAGNKAVKNVV-----SGLKKFGDANFP----- 617  
QY 217 FVAFSADVTPVNYKYARRTVNMGDEPS-----SRILANTNSITDVSWIYS---LAG 264  
Db 618 -LTSADNLTKQYDNAYKGLTNLDEKSGKQTPVADNTAATVCDLRLGLGWISADKTTG 676  
QY 265 TNTKYQSFNSNYGPGSTGYLYFPYKLVKAADANNVGLQYKLNNGVQVGFATSTSANNTT 324  
Db 677 ESKEYSAQVRNANE-----VFKSGNGINVSCKTLTDNGTREITELAKDEN--- 722  
QY 325 ANPTPAVDEIKVAKIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVAPMI 369  
Db 723 -----ATAFGSGKALRDNVTVAIGTGNVVAENAEKSGAGDPNIEDKA 764  
QY 370 GNYLSSNEN 379  
Db 765 GGSYAFGNDN 774

## RESULT 10

US-09-541-782-6  
; Sequence 6, Application US/09541782  
; Patent No. 6284480  
; GENERAL INFORMATION:

; APPLICANT: Nislow, Corey  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Antifungal Assay  
; FILE REFERENCE: 1015  
; CURRENT APPLICATION NUMBER: US/09/541,782  
; CURRENT FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1073  
; TYPE: PRT  
; ORGANISM: S.pombe  
US-09-541-782-6

Query Match 6.8%; Score 135.5; DB 4; Length 1073;  
Best Local Similarity 20.9%; Pred. No. 0.067;  
Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;

QY 14 NGO-TQLEAARMELTDLINAKA-MTLASL-----QDYAKIEASLSAYSEATVNNLNAT 67  
Db 611 NGYFTLLDNDFASMEELLNTHSNQLLISMTKITFHFQSLDEALQASCAVPNSSLDLI 670  
QY 68 LEQLMAKTNLESAINQANTDKTTFDNEHPN-----LVEAYKALKTT---LEQRATNLEG 119  
Db 671 VSELKDSKNSLLDALEHSLQDISMSQKLGNGISSLELIELQDKMKESYQLVQELRSLYN 730  
QY 120 LSTAYNQIRNNLVLYNKASSLIITKTLPLNGGTLDSNEIT-----TA 164  
Db 731 LQHTHEESQKELMYGVNRDIDALVKTCTTSLNDADILSDQSKPESKQOQLIANI 790  
QY 165 NKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEOSFVGTETFNANVQPSN 214  
Db 791 GKIVSNFLOEQNESLYTKADILHSLNDTNSNIRKANEIMNRSEEF---RNA----- 841  
QY 215 YSFVAFSADVTPVNYKYARRTVNMG-----DEPSSRLANTNSITDVSWIYSLA----- 263  
Db 842 -----ASQAEIVGANKERIOKTVENGSQLDSDSKSKAHSNRSMYD---HCLALAESQK 893  
QY 264 GNTKYQ-----FSSNYGPGSTGYLYFPYKLVKAADANNVGL---QYKLNNGVQ 310  
Db 894 GVNLEVQTLDRLLQKVKEHSEDNTEKHOOL---LDLLESLVGNNDNLIDSIKTPHTELQ 950  
QY 311 QVE-----FATSTSANNTT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 351  
Db 951 KITDHLVKGTGTSANTHNELLGDESCLNLETTIEDTSLVKLETGDTGTPSKRELPAFPS 1010  
QY 352 -----ELSVPTGEGNMKNVAPMIGNIYLSNENNADKIPGRRP 390  
Db 1011 WTRDSSLIKETTNLNDLSDKKFVREITYTSSNQTNPDV---YDKP 1052

## RESULT 11

US-09-268-347-30  
; Sequence 30, Application US/09268347  
; Patent No. 6335182  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena M.  
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
; FILE REFERENCE: 1038-860  
; CURRENT APPLICATION NUMBER: US/09/268,347  
; CURRENT FILING DATE: 1999-03-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 30  
; LENGTH: 1004  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-09-268-347-30

Query Match 6.7%; Score 134; DB 4; Length 1004;

Best Local Similarity 20.3%; Pred. No. 0.079;  
Matches 100; Conservative 53; Mismatches 160; Indels 180; Gaps 22;

QY 1 GCMSTTKDANPNNGOTOLEARM---ELDTLNKAMTASLQDYAKIEASLSAYSE 56  
Db 346 GKLAKTKVLKVSANGPNVKISNVADGENTDAVSFK--QKALQD-KQVTLASNAYAN 402  
QY 57 A-----ETVNNLNATLEQLKMAKTNLESAINQANTDKTTF-----DN-- 94  
Db 403 GGSADGCGKIQTLSNGLN-----PKFKSTDGELLNKAENDTVTFPPKGSVQVGDGK 457  
QY 95 -----EHPNLVEAYKAL-----KTTL--- 110  
Db 458 ATIQDGAKTTLGLVEASELVDLSNKLKGVGTGDTGVTGDTHTDTLVKGDKVTLKAG 517  
QY 111 -----EQRATN-----LEGLSTAYNOIRNNLVLYNKASSLITK---TLDPNGGT 154  
Db 518 DNLKVKQGGTFTYALKDELTDVKSVERKDTANGA-----NGASTKITDKGLTITPANGAG 573  
QY 155 LLDSEIITANKNINNTLSTINEQKTNADALSNSPIKKVIONNEQSFVGTFTNANVQPSN 214  
Db 574 AAGA-----NTANTISVTKDGISA--GNKAVKNV-----SGLKKFGDANFDP-- 614  
QY 215 YSFVAFSADVPVNVKYARRVWNGDEPS-----SRILANTNSITDVSWIYSLAGT 265  
Db 615 ---LTSSADNLTQYDNAVYKGLTNLDEKSGKQTPTVADNTAATVGLRGLGWISADKT 671  
QY 266 ---NTKYQFSFSGVPGTGYLYFPYKLVKAADANNVGLYKLNNGNVQVVEFATSTSAN 321  
Db 672 KGELNKEYNAQVRNANE-----VKFKSGNGINVSGLTDLGTRFITFELAKDEN 720  
QY 322 NTTANPTPAVDEIKVAKIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVA 366  
Db 721 -----AIAFGSGSKALRDNVTAIGTGNVYNAEKSAGFGDPNIE 759  
QY 367 PMIGNIYLSNEN 379  
Db 760 DRAGGSYAFGNDN 772

RESULT 12  
US-08-923-992A-2  
; Sequence 2, Application US/08923992A  
; Patent No. 6280738  
; GENERAL INFORMATION:  
; APPLICANT: Tai, Joseph Y.  
; APPLICANT: Blake, Milan S.  
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
; TITLE OF INVENTION: Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923,992A  
; FILING DATE: 05-SEP-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION NUMBER:  
; FILING DATE: 06-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1164 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-923-992A-2

Query Match 6.6%; Score 131.5; DB 4; Length 1164;  
Best Local Similarity 22.8%; Pred. No. 0.15;  
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;

QY 7 KKDANPNNGOTOLEA-ARMELTDLINA-KAMTLASIQDYAKIEASLSAYSE-----AET 59  
Db 230 RKQAOADKDEAEVKVREELGKLFSSTKAGLDQEIQEHVKKETSSEENTQKVDEHYANS 289  
QY 60 VNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 108  
Db 290 LONLAQKSLLELDKATTN-----EQATQVKNQFLENAOKLKEIQPLIKETNVKLYKAMSE 344  
QY 109 TLEQRATNLEGLSSPAYNOIRNNLVLYNKASSLITKTLDPNGGTLLDSNEITTTANKNI 168  
Db 345 SLEQVEKELKHNSA-----NLEDVAKSKEIVREYEGKLNQKNLP--ELKQLEEEA 395  
QY 169 NNTLSTINEQ-----KTNADALSNSFIKKVI---QNEQSFVGTFTNANVQPSNYSFVAF 220  
Db 396 HSKLKQVVEDFRKFKTSEQVTPKRVKRDLAANENNOOKI-----ELTVSPENI--- 445  
QY 221 SADVTPVNVKYARRVWNGDEPSRILANTNSITDVSWIYSLAGTNTKYQFSFS-----N 275  
Db 446 -----TVYEGEDVKFTVAKSDSKTILDFSDLL-----TKYNPSVSDRISIN 487  
QY 276 YGPSTG---YLYFPYKLVKAADANNVGLYKLNNGNVQVVEFATSTSANNTTANP-TPAV 331  
Db 488 YKTNNDNHKIAEITIKNLKLNESQVTVLKAQDSDGNVVEKFTTITVQKKEEKQVKTPEQ 547  
QY 332 DEIKVAKIV 340  
Db 548 KDSKTEEKV 556

RESULT 13  
US-08-923-992A-4  
; Sequence 4, Application US/08923992A  
; Patent No. 6280738  
; GENERAL INFORMATION:  
; APPLICANT: Tai, Joseph Y.  
; APPLICANT: Blake, Milan S.  
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
; TITLE OF INVENTION: Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923,992A  
; FILING DATE: 05-SEP-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/024,707

Query Match 6.4%; Score 129; DB 4; Length 2048;  
Best Local Similarity 21.2%; Pred. No. 0.49;  
Matches 83; Conservative 60; Mismatches 156; Indels 92; Gaps 19;

QY 5 ITKDKAN-----PNNQOTQLEAARMELTDLINAKAMTLASLODYAKIEASLSSAYSEAE 58  
Db 288 VTQSDNNSRPAYTPN---TQALDPKFOATNNTKAGPLSIGSNSIKRKI-INVAGVYKTD 343  
QY 59 TVNNNLNATLEQLKMAKTNLESAINQANTDKITFDNEHNLVEAYKALTTLEORATNLE 118  
Db 344 AVN-----VAQLEAVKWKERRITFOGD-DNSTDVKIGLDNTL-----TIK 384  
QY 119 GLSSTAYNOIRNNLYDYLNKA--SSLITKTLPLNGTLLDSNEITANKNNNTLSTIN 176  
Db 385 GGAET--NALTNNIGVVKREADNSGLVKLAKTLNLT-----EVTNTLNATTVVKVGS 437  
QY 177 EOKTNADALSNSFIKKVIONNEQSFVGTFTNANVPQSNYSFVAFSADVTPVNYKYARRTV 236  
Db 438 SSTTAELLSDSL-----TFTQPN----- 456  
QY 237 WNGDEPSSRLANTNSI---TDVSWIYSLAGTN--TKYQFSFNSYGPSTGYLYFPYKLVKA 292  
Db 457 -TGSQSTSTVYGVNGVKFTNNAETTAAGTTTRITRDKIGFARDG-DVDEKQAPYLDKKQ 514  
QY 293 ADANNVGLQYKLNNGNVQVEFATSTANNTANPTPAYDVEIKVAKIVL-SGLRFGQNTI 351  
Db 515 LKVGSAI--TIDNG-IDAGNKKISLNKAGSSANDAVTIEQLKAAKPTLNAGAGISVTPT 571  
QY 352 ELSVPTGEGNMKNKVPAMIGNIYSSNNENAD 382  
Db 572 EISVDKSGNV--TAPTY-NIGVKTTTELNSD 599

RESULT 15  
US-08-923-992A-6  
Sequence 6, Application US/08923992A  
Patent No. 6280738  
GENERAL INFORMATION:  
APPLICANT: Tai, Joseph Y.  
TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B  
TITLE OF INVENTION: Streptococcal Beta Antigens  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/923,992A  
FILING DATE: 03-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,707  
FILING DATE: 06-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1128 amino acids  
TYPE: amino acid

Query Match 6.5%; Score 131; DB 4; Length 1104;  
Best Local Similarity 20.2%; Pred. No. 0.15;  
Matches 85; Conservative 50; Mismatches 154; Indels 132; Gaps 17;

QY 7 KKDANPNNGTQLEA-ARMELTDLINA-KAMTLASLODYAKIEASLSSAYSEAE 59  
Db 199 RQQAQADKKEDAQVREELGKLFSTRAGLDQQIQEHVKKETSSEENTQKVDEHVANS 258  
QY 60 VNNNLNATLEQLKMAKTNLESAINQANTDKITFDNEHNL-----VEAYKALYT 108  
Db 259 LQNLAKQVLEELDQATN-----EQATQVKNQFLNNAQKLKEIQPLIKETNVLKRYKAMSE 313  
QY 109 TLEORATNLEGLSSTAYNOIRNNLYDYLNKASSLITKTLPLNGTLLDSNEITANKNI 168  
Db 314 SLEQVEKELHNSSEA-----NLQDLVAKSKEIVREYEGKLNQKNLP--ELKQLEEEA 364  
QY 169 NNTPLSTINEQ-----KTNADALSNSFIKKYI---QNNQSFVGTFTNANVPQSNYSFVAF 220  
Db 365 HSKLKQVVEHFRKFKETSEQVTPKRVKRDLANENNQOKI-----ELTVSPENI----- 414  
QY 221 SADVTVPVNYKYARRTVWNGDEPSSRLANTNSITDVSWSIYSLAGTITKYQFS--FSNYGP 278  
Db 415 -----TYEGEDVKFTYAKSDS-----KTTLDFSDLLTKYNP 447  
QY 279 STGYLYFPYKLVKAADANNVGLQYKLNNGNVQVEFATSTANNTANPTPAYDVEIKVAK 338  
Db 448 SV-----SDRISTNYKN-----TDNHKIAE 468  
QY 339 IVLSGLRFGQN-TIELSVPTGEGNM-----NKNVAPMIGNIYSSNNENADKI 384  
Db 469 ITIKNLKLNQSQVTLKAKDDSGNVVEKFTITVQKKEKQVPTPEQKHSKTEQNVPOE 528  
QY 385 P 385  
Db 529 P 529

RESULT 14  
US-09-268-347-48  
Sequence 48, Application US/09268347  
Patent No. 6335182  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena M.  
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
FILE REFERENCE: 1038-860  
CURRENT APPLICATION NUMBER: US/09/268,347  
CURRENT FILING DATE: 1999-03-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 48  
LENGTH: 2048  
TYPE: PRT  
ORGANISM: Haemophilus influenzae  
US-09-268-347-48

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-923-992A-6

Query Match 6.3%; Score 126.5; DB 4; Length 1128;  
Best local Similarity 22.5%; Pred. No. 0.33;  
Matches 83; Conservative 51; Mismatches 158; Indels 77; Gaps 15;

QY 7 KKDANPNNGOTOLEA-ARMELTDLINA-KAMTLASLODYAKIEASLSAYSE-----AET 59  
Db 194 RKQAQADKKDEAEVKVREELGKLFSTKAGLDQEIHEHVKKETSSSEENTOKVDEHYANS 253  
QY 60 VNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 108  
Db 254 LONLAQKSLEELDRAATN-----EQATQVKNQFLENQAKLKEMQPLIKETNVKLYKAMSE 308  
QY 109 TLEQRATNLEGLSTAYNQIRNNLVLYNKASSLITKTLDPNGGTLDDSNEITANKNI 168  
Db 309 SLEQVEKELKHNSA-----NLEDLVAKSEIVREYEGKLNQSKNLP--ELKQLEEEA 359  
QY 169 NNTLSTINEQ-----KTNADALSNSFTKKVI---ONNEQSFVGFTNANQPSNYSFVAF 220  
Db 360 HSKLKQVVEDEPRKKFKTSEQVTPKRVKRDLAANENQOKI-----ELTVSPENI 409  
QY 221 SADVTPVNYKYARTVWNGDEPSSRIILANTNSITDVSWIYSLAGTNTKYQSFSS-----N 275  
Db 410 -----TVYEGEDVKFTVTAKSDSKTTLDFSDLL-----TKYNPSVSDRISTN 451  
QY 276 YGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGNQVQVEFATSTANNTANP-TPAV 331  
Db 452 YKTNTDNHKIAEITIKNLKLNESQVTILKAKDDSGNVVKEFTTITVQKKEEKQVPKTPEQ 511  
QY 332 DEIKVAKIV 340  
Db 512 KDSKTEEKV 520

Search completed: June 12, 2002, 10:49:10  
Job time: 203 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:49:47 ; Search time 29.99 Seconds  
(without alignments)  
1262.394 Million cell updates/sec

Title: US-09-147-052-4\_COPY\_693\_1086

Perfect score: 2004

Sequence: 1 GCMSTTKDANPNNGQTLE.....SSNENNADKIPGYRRPGTFL 394

-Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

-Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	831.5	41.5	702	2	S48754	major surface prot
2	754	37.6	702	2	S48753	major surface prot
3	744	37.1	650	2	S48751	major surface prot
4	743	37.1	649	2	S48752	major surface prot
5	713	35.6	647	2	A49218	hemagglutinin homo
6	575.5	28.7	386	2	S48755	major surface prot
7	189.5	9.5	320	2	S51560	major surface prot
8	175	8.7	6713	2	B99921	hypothetical prote
9	173	8.6	1302	1	J66009	surface-located me
10	165.5	8.3	661	2	A24222	hypothetical prote
11	164.5	8.2	2481	2	D90011	FmtB protein [impo
12	163.5	8.2	3890	2	C89921	hypothetical prote
13	159	7.9	933	2	S41539	fibrinogen-binding
14	156	7.8	807	2	B71605	hypothetical prote
15	155	7.7	1072	2	A68827	hypothetical prote
16	155	7.7	4688	2	F82885	hypothetical prote
17	154.5	7.7	682	2	S44131	subtilisin-like pr
18	154.5	7.7	1365	2	T30822	lmp1 protein - Myc
19	151.5	7.6	1237	2	D71850	probable outer mem
20	151	7.5	624	2	PC6003	surface membrane p
21	150.5	7.5	1051	2	T18351	lmp1 protein - Myc
22	145	7.2	2399	2	D71879	toxin-like outer m
23	144.5	7.2	1487	2	AG2560	hypothetical prote
24	144	7.2	1107	2	AC0976	probable autotrans
25	143.5	7.2	751	2	T40462	ser-lys rich hypot
26	142	7.1	135	2	B49218	hemagglutinin homo
27	142	7.1	2660	2	E95822	probable invasiv z
28	141.5	7.1	1645	2	F36907	phage-related prot
29	140.5	7.0	3194	2	D71917	toxin-like outer m

30 139.5 7.0 2401 2 T28676 rhoptry protein -  
31 138.5 6.9 1314 1 TNBYR6 transcription regu  
32 137.5 6.9 820 2 T17519 cell surface antiq  
33 137.5 6.9 926 2 AE1130 conserved hypothet  
34 137.5 6.9 989 2 D89852 fibrinogen-binding  
35 137.5 6.9 1524 2 S68553 surface layer prot  
36 137.5 6.9 4152 2 T31102 filamentous hemagg  
37 137 6.8 719 2 S55119 hypothetical prote  
38 137 6.8 5005 2 F82884 hypothetical prote  
39 136.5 6.8 568 2 E97066 membrane associate  
40 136.5 6.8 1073 2 S14032 kinesin-related pr  
41 136.5 6.8 2271 2 F90073 hypothetical prote  
42 135.5 6.8 1085 2 T38378 kinesin-like prote  
43 135.5 6.8 1238 2 A64596 hypothetical prote  
44 135 6.7 2269 2 T28677 rhoptry protein -  
45 134.5 6.7 589 2 B97806 hypothetical prote

ALIGNMENTS

RESULT 1

S48754

major surface protein (clone pMGAL.4) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: S48754

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.  
FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739

A:Accession: S48754

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAR>

A:Cross-references: EMBL:L28424

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 41.5%; Score 831.5; DB 2: Length 702;  
Best Local Similarity 42.0%; Pred. No. 3.6e-38;  
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY 9 DANPNNG-----QTQLEAARMELTDLINAKAMFTLASLDYAKIEASLSAYSEATVNN 62  
DB 68 NTPNGGGTDAQAQQAALAAKKELSDLLATQNSLSTYADYANIQTTLTAAYTTAKSTD 127  
QY 63 NLNATLEQLKMAKTNLKESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSS 122  
DB 128 NTSATLEQVKTSATSLQTAIDTAASKTSFDEKNPELIKAYNAKKTLLKWRNLSGLTD 187  
QY 123 TAYNQRNLDVLYNKASLLITKTLDPLNGGTLLDSNEITTANKNNNTLSTINEQKTN 182  
DB 188 SNFATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAVSQANTNLSNAVKLETWKTN 246  
QY 183 DALNSFIKKVTONNEQSFVGTFTNANVOPSNYSVFASFADVTP-----VNYKYARRTV 236  
DB 247 TVLATSFVKVEVLVKNKLTGIDT-TNNQEPGNYSFVGYSVDVTTGSDNARNPNWSFAQRKV 305  
QY 237 WNGD-----EPSSRILANTNSITDVSWIYSLAGTNTKYQFSFSNYPGSTGYLYFPYKLVK 291  
DB 306 WTSNTDILSQPOPAECENQOSAPDVSWIYNLTGGMGAKYSLTFNYGYPGSTGFLYFPKLVN 365  
QY 292 AADANNVGLQYKLNNGNVQOQEFATS-----TSANN-----TTANPTPAYD 332  
DB 366 SSDSDKVALEYKLNESAVKTIIDFSPSQTSPVASDASTRENNRSTAAPAQGSTINPAPTLD 425  
QY 333 EKKVAKIVLSGLRFGQNTTSLVSP-TGEGNMKVAPMIGNIYLSNENNAADK 383  
DB 426 DIKIAKVTLSNLKFGNTSTIEFSVPTTAKGTSKVAPMIGNIMKLTSSDRDVK 477

RESULT 2

S48753  
major surface protein (clone pMGAL.3) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: S48753  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;  
FEBS Lett. 352, 347-352, 1994  
A:Title: The organisation of the multigene family which encodes the major cell surface  
A:Reference number: S48751; MUID:95010739  
A:Accession: S48753  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-702 <MAR>  
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62417.1; PID:g535690  
C:Genetics:  
A:Genetic code: SGC3  
A:Start codon: GTG

Query Match 37.6%; Score 754; DB 2; Length 702;  
Best Local Similarity 41.6%; Pred. No. 6.2e-34;  
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;  
QY 11 NPNNGOT-----OLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVNNLNA 66  
Db 69 NPNNGTPEQQLAARKLTLDLCTENTNVALYADYAKIQSTLSAYMTAKTASNTSA 128  
QY 67 TLEQLKMAKTNLESAINOANTDKTTFDNEHNLVAYKALKTTLEQRTNLEGLSSSTAYN 126  
Db 129 TLENLRASSTTLQAAIDKAANDKRVDSVNOPLVAAAYNNLTKTSKTSLSGLSENKYG 188  
QY 127 QIRNNLVLYNKASSLTITKPLNG--GTLDSNEITITANKNINNTLS--TINEQKNTNA 182  
Db 189 GIKNHLSKLFDTGSAITAKTLDPTSGERTLEKVN---ANNIKMAISPESLKKWKGA 245  
QY 183 DALNSFIKKVIONNQSFVGTFTNANQPSNYSFVAFSADVTP-----VNYKYARTVW 237  
Db 246 DKF-NEFEKNPLSKELKSTSDTAHQEQPANWFAAYSVDLTSNQNPNNFNAQRKYW 304  
QY 238 NGB-EPSSRILANTN-SITDYSWISYSLAGTNTKYQFSNKGPSGTYLYFPYKLVKAAD 294  
Db 305 TSENQPGKATLVSSPSVATDVSWSIYSLAGTGKTYLTETFGPDNAPFLYLPYKLVKAAD 364  
QY 295 ANNVGLQYKLNNGVQOQVEP-----ATST-----SAN 321  
Db 365 SSSVALQYSLNKTSSKLNFKPAETVSTNTDQSENEVATTSTTEARSSYKVLVADEAATS 424  
QY 322 NTTANTPPAYDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSSENNA 381  
Db 425 NNEHNPTVSDINIAKVTLSGLTGTGNTIEFSVPEG-----KVAPMIGNIYLSSESQ 479  
QY 382 DK1 384  
Db 480 VK1 482

RESULT 3

S48751  
major surface protein (clone pMGAL.1) precursor - Mycoplasma gallisepticum  
N:Alternate names: major hemagglutinin pMGAL  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: S48751; A44793  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;  
FEBS Lett. 352, 347-352, 1994  
A:Title: The organisation of the multigene family which encodes the major cell surface  
A:Reference number: S48751; MUID:95010739  
A:Accession: S48751  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-650 <MAR>

A:Cross-references: GB:U90714; EMBL:L28423; NID:g1905869; PIDN:AAB50152.1; PID:g19058  
R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whithear, K.G.  
Infect. Immun. 60, 3885-3891, 1992  
A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepti  
A:Reference number: A44793; MUID:92363591  
A:Contents: S6  
A:Accession: A44793  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 26-42 <MAR>  
A:Note: sequence extracted from NCBI back bone (NCBIP:111017)  
C:Genetics:  
A:Genetic code: SGC3  
A:Start codon: GTG

Query Match 37.1%; Score 744; DB 2; Length 650;  
Best Local Similarity 42.9%; Pred. No. 2e-33;  
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;  
QY 11 NPNNGO-----TOLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVNNN 63  
Db 51 NPGDGGGMNNAQAQELAAARMGLTTFVDSRAKNLGLYDYKKTQNTLTAKYDAKTVLDN 110  
QY 64 LNATLEQLKMAKTNLESAINOANTDKTTFDNEHNLVAYKALKTTLEQRTNLEGLSST 123  
Db 111 SSSITQNLNEAKTRLETAIRTAATSKOTFDEQHAELVKVYKELKTTLSNETPATLPTADA 170  
QY 124 AYNQIRNNLVLYNKASSLTITKPLNG--GTLDSNEITITANKNINNTL--STINEQKTN 181  
Db 171 QYAGIKMHLSDVGLGDIAGKAITTKTLEPVGDP--LTAGAVTMANTKIVEAIKDEVLNPKKEN 229  
QY 182 ADALNSFIKKVIONNQSFVGTFTNANQPSNYSFVAFSADVTP-----NYKYAR 233  
Db 230 ATKLADSFVKVLYKEKITGVEEAHN--KAQPAHYSEFVGYSDVITGTANGQTSIPNNYQAQ 288  
QY 234 RPVW-NGDEPPSRILANT-----NSITDYSWISYSLAGTNTKYQFSNKGPSGTYLY 284  
Db 289 RTIFTNGDEP--RSYSNTPVDGQTMQAQPLSNVSVWISYSLAGTGAKTYLFTFYGPSTGYLY 346  
QY 285 FPKYKLVKADANNVGLQYKLNNGVQOQVEFATSTSANNTAN--PTPAVDEIKVAKIVLSG 343  
Db 347 FPKYKLVNTSDQMKLGLEYKLND-----TEPSAITFEGNEQTMNGKTPVNDINNAKVTLAN 402  
QY 344 LRFQONTIELSVPTGEGNNKVPAMIGNIYLSSENNAADKI 384  
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 438

RESULT 4

S48752  
major surface protein (clone pMGAL.2) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: S48752  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.  
FEBS Lett. 352, 347-352, 1994  
A:Title: The organisation of the multigene family which encodes the major cell surface  
A:Reference number: S48751; MUID:95010739  
A:Accession: S48752  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-649 <MAR>  
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62416.1; PID:g535689  
C:Genetics:  
A:Genetic code: SGC3  
A:Start codon: GTG

Query Match 37.1%; Score 743; DB 2; Length 649;  
Best Local Similarity 42.3%; Pred. No. 2.2e-33;  
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

Qy 9 DANPNNGO-----TQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSSEATVN 61  
 Db 48 DTNPGDGGGMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTAKYDAAKTVL 107  
 Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLEVEAYKALKTTLEQRATNLEGLS 121  
 Db 108 DNSSSTTQNLNEAKTRLETAIRTAATSQTDFEQAELVYKELKTKTTLSETATLAPYA 167  
 Qy 122 STAYNOIRNNLVLYNKASSLITKTDPLNGTLLDSNEITTANKNINNTL--STINEQK 179  
 Db 168 DAQYAGIKMHLISGLYDAGKAITTKTLEPEVGDPLTASAVMANTKIIVEAIKDEVINPQK 226  
 Qy 180 TNADALSNSFKKVIQNNQESFVGTFTNANVQPSNYSFVAFSADVPV-----NYKY 231  
 Db 227 ENATKLADSFVKQVLVKEKITGVVEAHN-KAQPANYSFVGSVDITGTTGTQTSIPNWDY 285  
 Qy 232 ARRTVW-NGDEPSSRIANT-----NSITDVSWIYSLAGTNTKYQFSFNSGVPSTGY 282  
 Db 286 AORTIFTNSDEP--RSISNTPADQOTMAQPLSNVSWIYSLAGTGAKYTLEFTYGPSTGY 343  
 Qy 283 LYFPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSA-----NNTTANPTPAVDEIKVA 337  
 Db 344 LYFPYKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDOTMNGKTPVNDINVA 395  
 Qy 338 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSSENENNADKI 384  
 Db 396 KVTLANLNFGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437

RESULT 5  
 A49218  
 hemagglutinin homolog pmGAL.2 - Mycoplasma gallisepticum  
 C:Species: Mycoplasma gallisepticum  
 C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 07-Dec-1999  
 C:Accession: A49218  
 R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.  
 Infect. Immun. 61, 903-909, 1993  
 A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglutinin  
 A:Reference number: A49218; MUID:93162830  
 A:Accession: A49218  
 A:Status: preliminary  
 A:Molecule type: DNA; protein  
 A:Residues: 1-647 <MAR>  
 A:Cross-references: GB:952116; NID:g265625; PIDN:AAB25397.1; PID:g265626  
 A:Experimental source: S6  
 A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBI:P:125183)  
 C:Genetics:  
 A:Genetic code: SGC3

Query Match 35, 68; Score 713; DB 2; Length 647;  
 Best Local Similarity 41.18; Pred. No. 9.6e-32;  
 Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

Qy 9 DANPNNGO-----TQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSSEATVN 61  
 Db 48 DTNPGDGGGMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTAKYDAAKTVL 107  
 Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLEVEAYKALKTTLEQRATNLEGLS 121  
 Db 108 DNSSSTTQNLNEAKTRLETAIRTAATSQTDFEQAELVYKELKTKTTLSETATLAPYA 167  
 Qy 122 STAYNOIRNNLVLYNKASSLITKTDPLNGTLLDSNEITTANKNINNTL--STINEQK 179  
 Db 168 DAQYAGIKMHLISGLYDAGKAITTKTLEPEVGDPLTASAVMANTKIIVEAIKDEVINPQK 226  
 Qy 180 TN-----ADALSNSFKKVIQNNQESFVGTFTNANVQPSNYSFVAFSADVPV----- 227  
 Db 227 ENATKLADSLSSIVKKITGVVE-----AHNKAQPANYSFVGYKRYTTELLDKQVFP 279  
 Qy 228 NYKYARRTVW-NGDEPSSRIANT-----NSITDVSWIYSLAGTNTKYQFSFNSGVP 278  
 Db 280 NWDYQAORTIFTNSDEP--RSISNTPADQOTMAQPLSNVSWIYSLAGTGAKYTLEFTYGP 337

Qy 279 STGYLYFPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSA-----NNTTANPTPAVDE 333  
 Db 338 STGYLYFPYKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDOTMNGKTPVND 389  
 Qy 334 IKVAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSSENENNADKI 384  
 Db 390 INVAKVTLANLNFGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 435

RESULT 6  
 S48755  
 major surface protein (clone pmGAL.5) precursor - Mycoplasma gallisepticum (fragment)  
 C:Species: Mycoplasma gallisepticum  
 C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
 C:Accession: S48755  
 R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.  
 FEBS Lett. 352, 347-352, 1994  
 A:Title: The organisation of the multigene family which encodes the major cell surface  
 A:Reference number: S48751; MUID:95010739  
 A:Accession: S48755  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-386 <MAR>  
 A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62419.1; PID:g535692  
 A:Note: the sequence of residues 385-386 and the corresponding nucleotide sequence ar  
 C:Genetics:  
 A:Genetic code: SGC3  
 A:Start codon: GTG

Query Match 28.78; Score 575.5; DB 2; Length 386;  
 Best Local Similarity 42.38; Pred. No. 1.6e-24;  
 Matches 138; Conservative 50; Mismatches 113; Indels 25; Gaps 10;

Qy 9 DANPNNG-----QTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSSEATVN 62  
 Db 64 NTNPGCGGMDNSAQALAAAKKELSDLLATQNSLSTVADYAKIKNDLTAAYTTAETASQ 123  
 Qy 63 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLEVEAYKALKTTLEQRATNLEGLS 122  
 Db 124 NQAATLEQVKNAASTLQTAINTAVNEKKYFDENNELSETAYTNLTKTLEGENTLAAAFND 183  
 Qy 123 TA-YNOIRNNLVLYNKASSLITKTDPLN--GTLLDSNEITTANKNINNTL--STINEQ 178  
 Db 184 SANYGGIKTHLLSLNQAKTITTTSL--LNDAGQSPKNDVVKINKEITDAINPILLNQO 241  
 Qy 179 KTNADALSNSFKKVIQNNQESFVG---TFTNANVQPSNYSFVAFSADVP-----VNY 229  
 Db 242 KANADMLATSFTKQVL-NDAQLTSGSSETSMQTQPGNYSFVGSVDVTGSSNNARPNN 300  
 Qy 230 KYARTVWNGDEPSSRIANTNSITDVSWIYSLAGTNTKYQFSFNSGVP--TCYLYFPYK 288  
 Db 301 NFAQRKVDVNTRAPLAQTEQSNKLDVSWIYSLSGMGAKYTVTFDYYGASNNAYLYFPYK 360  
 Qy 289 LVKAADANNVGLQYKLNNGNVQVFE 314  
 Db 361 LVQVND--NVGLQYVNNTPPKLVNF 384

RESULT 7  
 S51560  
 major surface protein (clone pmGAL.6) - Mycoplasma gallisepticum (fragment)  
 C:Species: Mycoplasma gallisepticum  
 C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
 C:Accession: S51560; S48757  
 R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.  
 FEBS Lett. 352, 347-352, 1994  
 A:Title: The organisation of the multigene family which encodes the major cell surface  
 A:Reference number: S48751; MUID:95010739  
 A:Accession: S51560  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA

[illegible]

```

QY 9 DANPNNGOTOLE-----AARMEITDLINAKAMTILASLDQYAKIEASLSSAYSEAE 58
Db 1186 DOAKTTGETSIDQVTPVTKKATARNETAILNNKLEIQATPDATDEEKQAADA--EAN 1243
QY 59 TVNNLN-----ATLEOLKMAKTWLESAINQANTDKTTFDEHPNLVYAYKALKTTL-- 110
Db 1244 TENGKANQAIISAATTNAQVDEAKANAAAIN-AVTPKVVKKQAANDEIDOLOATQNTNVIN 1302

```

Db 1360 K SNAKNEVDQAVTTONQAIDNTTGATTEEKNAAKDL-----VLKAKEKAYQDIL---N 1409

Db 1410 AQTNN-----DVTQI-----KDQAVADIQGITA DTTIKDVAKDELAT 1446

OV 266 NTKYQFSFSNYGPGSTGYLYFFPYKLVKAADAN-----NVGLQYKLNNGNVQQVEFATST 318

Db 1447 KANEQKAL-----IAQTADATTEEKEQANQQVDAQLTQGN-QNIENAQSI 1490

OV 319 SANNTTA-NPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMNKVAPMIGNIYLSN 377

Db 1491 DDVNTAKDNAIOAIDPIOASTDVKTNR-----AELLTEM-----ONKITEILNNNETTNE 1541

Ov 378 ENNAD 382

Db 1542 EKGND 1546

RESUL  
12  
C89921

C:Species: *Staphylococcus aureus*

C;Date: 10 MAY 2001 \*sequence\_revision to may 2001 \*change to 000 2001  
C;Accession: C89921

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki

Lancet 357, 1225-1240, 2001

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Status: preliminary

A;Residues: 1-3890 <KUR>

A; Experimental source: strain N315

A; Gene: ebhB

Best Local Similarity 24.4%; Pred. No. 1.1;

**TITLE**

— — — — —

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

*[The following text is extremely faint and appears to be bleed-through from the reverse side of the page. It is largely illegible.]*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Db 3637 -----INSGTTVAGVHTVQSNANTLDQAMNTLRQSIANNDATKASEDYVDA 3682

QY 191 ---KKVIONNEQSFVGTFTNANVQPS-NYSFVAFSADVTPVNYKYARRTVWNGDE----- 241

[illegible]

Db 3683 NNDKQTAYNNAVAEAETIINANSNPNPSTITQKAE--QVN---SSKALNGDENLATA 3737

QY 242 -PSSRLANT-NSITD-----VSHYIS---LACNTKTYQFSFSGYNGPSTGYLYFPYKLV 290

Db 3738 KQAKTYLNTLTSITDAOKNNLISQISSATRVSGVDTVKQNA-----QHL 3782

QY 291 KAADANNVGLQKLNNGVQVFEFATSSANTTANPTPAVDE-ISKVAKIVL---SGLRF 346

Db 3783 DQAMAN-----LONGINNESQVSKSEKRYDADTNKQOEYNATKAAILNKSTGPNT 3835

QY 347 GONTIELS---VPTGEGNNKVPAMI 369

Db 3836 AQNAVENALQRYNTAKDALNGDAKLI 3861

RESULT 13

S41539

fibrinogen-binding protein - Staphylococcus aureus

N:Alternate names: clumping factor

C:Species: Staphylococcus aureus

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Oct-1999

C:Accession: S41539; S36630

R:McDevitt, D.; Francois, P.; Vaudeaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus S41539

A:Reference number: S41539; MUID:94224142

A:Accession: S41539

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-933 <MCD>

A:Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

Query Match 7.9%; Score 159; DB 2; Length 933;

Best Local Similarity 22.8%; Pred. No. 0.3;

Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

QY 3 MSITKDPNNGQTL--EARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEIV 60

Db 73 VSDTKTSNTNGETSVAGNPAQOETQSSSTNATT-----ETPTVTEATT 120

QY 61 NNLANLTLQKMAKTNLESAINQANTKTTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120

Db 121 TTNQANTPATTOSSNTNAELVQ-TSNETTND-TNTVSSVNS-----PQNSINAEV 172

QY 121 SST-----AYNQIRNLDVLYNK-----ASSLITKTLPLNGGTL 156

Db 173 STQDSTSTPATPSNNEAPQSTDAKNDVYVNAVNTSAPRMAFSLAAVAADAPAAAGTDI 232

QY 157 DSNEITTANKNTNNTLTINEQ----KTN-ADALNSFTK---KVIONNEQSFVGTFTN 207

Db 233 -TNQLTNVTGIDSGTVYPHQAGYVKLYGFSVPNSAVKGDFTKITVPKELNLCVST 291

QY 208 ANQPSNYSFVAFSADVTPVNTKYARRTVWNGDEPSSRLANTNSTIDVSYWISLAG-TN 266

Db 292 AKVPP-----IMAGDQ---VLANGVIDSDGNVIYTFDYV 324

QY 267 TKYQFSFSGYNGPSTGYLYFPYKLVKAADANNVGLQKLNNGVQVFEFATSSANNITAN 326

Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362

QY 327 PPAVDEIKVAKI-----VLSGLRFGONTIELSYPTGEGNNKVPAMI-GNIYLSNNE 378

Db 363 KTVLDVDEYKGFYNLISIKGTIDQIDKTNTRYQTIIYVNPNGDNVIAPIVLTGNLKPNTDS 422

QY 379 N 379

Db 423 N 423

RESULT 14

B71605

hypothetical protein PFB0850c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000

C:Accession: B71605

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Portea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H. Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600; MUID:99021743

A:Accession: B71605

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-807 <GAR>

A:Cross-references: GB:AE001420; GB:AE001362; NID:g3845287; PIDN:AAC71959.1; PID:g384

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0850c

Query Match 7.8%; Score 156; DB 2; Length 807;

Best Local Similarity 23.8%; Pred. No. 0.36;

Matches 73; Conservative 48; Mismatches 124; Indels 62; Gaps 14;

QY 60 VNNLNATLEQLKMAKT-----NLESAINQA--NTDK-----TTFDNEHPNLVEAYKALKT 108

Db 468 INNLSYT--QLKMNINFINHESPINQHNNTFKVNDTNEFEPTNKKKKKKKK 525

QY 109 TLEQRATNLEGLSSTAYNOIRNLDVLYNKASLITKTLPLDPLNGGTLTDSNEITTANKMI 168

Db 526 NTHFNNNNNNNKCLYDINQ--DHNSINTNQFDHI-----NNVKTEQNL 574

QY 169 ---NNTLSTINQKTNADALSNSFTKVIQNEQSFVGTFTNANQPSNYSFVAFSADVT 225

Db 575 QKKHNMKSQVSKSNKNNKNSHLKQININ-----TNNMDKNNSHISKVIVD 626

QY 226 PVNYK-----YARRTVWNGDEPSS---RILANTNSITDVSWIYSLAGNTIYQ-PSFSN 275

Db 627 DNKLSSHADNSNEIVTKGKKKNTNKKKINNINSVNNVNNINSNNNNIISMNN 686

QY 276 YGPSTGYLYFPYKLVKAADANNVGLQYK-----LNNGNVQVQEFATSTAN 321

Db 687 VNNMNNMYPFVNIQ-KDDSNIALLYNNKPNIDFNNFOLNHNHNTQNNNTNNVMLN 745

QY 322 N--TTAN 326

Db 746 NNLITSN 752

RESULT 15

A86827

hypothetical protein yqfG [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C:Accession: A86827

R:Boletín, A.; Wincker, P.; Mauger, S.; Jallón, O.; Malarne, K.; Weissensbach, J.; Eh Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: A86827

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1072 <STO>

A:Cross-references: GB:AE005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yqfG

Query Match 7.7%; Score 155; DB 2; Length 1072;

Best Local Similarity 21.0%; Pred. No. 0.6;

Matches 97; Conservative 76; Mismatches 154; Indels 136; Gaps 20;

QY 9 DANPNNGQTGLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSE-AETVNNNL--- 64

```

Db 430 ESNVSQDTSSEAS-TNSNGSISLSPSNISSTSD--SESATNSSDFSNVAEVANNLSASV 486
QY 65 -NATLEOLKMAKTNLESAINQANTDKTTFONEHNLVYAYKALKTITLEQRATN----- 116
Db 487 NNSSSVLSSTSTADNLGINOSGSDNLTKDSSEISTSGAFLSSNOTSSEASTNSNSISL 546
QY 117 -----LEG-LSSTAYNQI----RNNLVDLYNKASSLI--TKTLDPL-----NGG 153
Db 547 SPSNISSTVLESTSSSNFNSVAEVANNLSASVYNNSSSVLSSTSTADNLEINQFGSDN 606
QY 154 TLLDSNEITTANKNINNTLSTINEQKNADALSNSFTKKVYQNNNEQSFVGTFTNANVOPS 213
Db 607 LTKDSEISTSG-----AFLSSNOTSSEASSNS---MSSINSPSLSLSTNSSESAT 655
QY 214 NYSFVAFSADVTPVNYKYARTVNGDEPSSRII-----ANTNSITDVSWIYSLAG 264
Db 656 NQS---NSSEATKYDN-----NSSTHSSNIISSGNSDSDSDSDSSNLSLSSPN 704
QY 265 TNTKYQFSFSGNYGPGTGYLPYKLVKAADANNYGLOYKLNGNVQOVFEFATSTSANNT- 323
Db 705 LETNOTIS-----SKPSEVNNI-----SENPKVSSNSVQENSTD 740
QY 324 ---TANTPTPAVDE-----IKVAKIVLGLRPFQONTIE---LSVPTGEG 360
Db 741 HEMSTNPKSSISPISTTSSSQKESQSNLLTTTEGINNPITFNSSSENSEAAASILATSYS 800
QY 361 MNKVAPMIGNIYLS-----SNNENADKI 384
Db 801 NNSSESSETGCLYISNEAQRDNGSEISHSLPSSNSNENNVSSI 843

```

Search completed: June 12, 2002, 10:49:50  
Job time: 208 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:51:05 ; Search time 17.44 Seconds  
(without alignments)  
874.742 Million cell updates/sec

Title: US-09-147-052-4\_COPY\_693\_1086

Perfect score: 2004

Sequence: 1.GCMSTTKDANPNNGQTGLE.....SSNENADKIPGYRRGTGL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

\*Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	154.5	7.7	682	1 NISP_LACLA	Q07596 lactococcus
2	143	7.1	857	1 AR56-CANAL	P78586 c arg5,6 pr
3	138.5	6.9	1314	1 SWI1_YEAST	P09547 saccharomyc
4	137	6.8	719	1 YM41_YEAST	Q03213 saccharomyc
5	135.5	6.8	1085	1 CUT7_SCHPO	P24339 schizosacch
6	132	6.6	1271	1 Y338_MYCGE	P47580 mycoplasma
7	131.5	6.6	1164	1 BAG_STRAG	P27951 streptococc
8	129	6.4	1251	1 RBP2_PLAVB	Q00799 plasmodium
9	128.5	6.4	2334	1 WAPA_BAGSU	Q07833 bacillus su
10	128.5	6.4	3712	1 LMA_DROME	Q00174 drosophila
11	128	6.4	995	1 YI09_YEAST	P40442 saccharomyc
12	127	6.3	956	1 YEF3_YEAST	P32618 saccharomyc
13	126.5	6.3	444	1 PST1_YEAST	Q12355 saccharomyc
14	126.5	6.3	1630	1 MSP1_PLAFK	P04932 plasmodium
15	126.5	6.3	1639	1 MSP1_PLAFW	P04933 plasmodium
16	125.5	6.3	1858	1 P3K2_DICDI	P54674 dictyosteli
17	124.5	6.2	821	1 LIN5_CABEL	P45970 caenorhabdi
18	124.5	6.2	1010	1 YK1_CABEL	P34278 caenorhabdi
19	124.5	6.2	1018	1 FNBA_STAAU	P14738 staphylococ
20	124	6.2	749	1 YK51_YEAST	P04957 saccharomyc
21	124	6.2	750	1 YK51_YEAST	P34231 saccharomyc
22	124	6.2	823	1 NSP1_YEAST	P14907 saccharomyc
23	124	6.2	865	1 MYSP_SCHJA	Q05870 schistosoma
24	124	6.2	1117	1 YN96_YEAST	P33753 saccharomyc
25	124	6.2	1159	1 N124_SCHPO	Q09904 schizosacch
26	124	6.2	1433	1 CAT8_YEAST	P39113 saccharomyc
27	123.5	6.2	537	1 ARP_PLAFA	P04931 plasmodium
28	123.5	6.2	1902	1 P2P_LACLC	P15293 lactococcus
29	123	6.1	1260	1 ALS1_CANAL	P46590 candida alb
30	123	6.1	1637	1 MKSP_STAAU	P80544 staphylococ
31	123	6.1	1790	1 USO1_YEAST	P25386 saccharomyc
32	122.5	6.1	507	1 FLIC_SALON	Q06974 salmonella
33	122	6.1	650	1 PDR4_YEAST	P19880 saccharomyc

## RESULT 1

NISP\_LACLA

ID NISP\_LACLA STANDARD; PRT; 682 AA.

AC Q07596;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nisin leader peptide processing serine protease NISP precursor

DE (EC 3.4.21.-).

GN NISP.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Lactococcus.

OX NCBI\_TaxID=1360;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NIZO R5;

RX MEDLINE=93239683; PubMed=8478324;

RA van der Meer J.R., Polman J., Beertuyzen M.M., Siezen R.J.,

RA Kuipers O.P., de Vos W.M.;

RT "Characterization of the Lactococcus lactis nisin A operon genes

RT nisp, encoding a subtilisin-like serine protease involved in

RT precursor processing, and nispR, encoding a regulatory protein

RT involved in nisin biosynthesis.";

RT J. Bacteriol. 175:2578-2588(1993).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=6F3;

RX MEDLINE=94213458; PubMed=8161176;

RA Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,

RA Hammelmann M., Entian K.-D.;

RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis

RT 6F3.";

RT Appl. Environ. Microbiol. 60:814-825(1994).

RN [3]

RP 3D-STRUCTURE MODELING.

RX MEDLINE=95357326; PubMed=7630881;

RA Siezen R.J., Rollemans H.S., Kuipers O.P., de Vos W.M.;

RA "Homology modelling of the Lactococcus lactis leader peptidase Nisp

RT and its interaction with the precursor of the lantibiotic nisin.";

RT Protein Eng. 8:117-125(1995).

CC -1- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.

CC -1- PATHWAY: LAST STEP OF NISIN BIOSYNTHESIS.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL

CC (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE

CC SUBTILASE FAMILY.

CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS

CC IN THE REGION OF THE MEMBRANE ANCHOR.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

or send an email to license@isb-sib.ch).
-----
EMBL: L11061; AAA25200.1; -
EMBL: X76884; CAA54210.1; -
HSP: P29600; IGI.
MEROPS: S08.059; -
InterPro: IPR001899; Gram_pos_anchor.
InterPro: IPR002029; Peptidase_S8.
Pfam: PF00082; Peptidase_S8; 1.
PRINTS: PR00723; SUBTILISIN.
PROSITE: PS00136; SUBTILASE_ASP; 1.
PROSITE: PS00137; SUBTILASE_HIS; 1.
PROSITE: PS00138; SUBTILASE_SER; FALSE_NEG.
PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KW Hydrolase; Serine protease; Cell wall; Signal; Transmembrane; Zymogen.
FT SIGNAL 1 22
FT PROPEP 23 195
FT CHAIN 196 682
FT DOMAIN 196 651
FT TRANSMEM 652 679
FT DOMAIN 680 682
FT ACT_SITE 259 306
FT ACT_SITE 306 312
FT ACT_SITE 512 512
FT DOMAIN 652 657
FT CONFLICT 500 500
FT SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;

Query Match 7.7%; Score 154.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.11;
Matches 75; Conservative 61; Mismatches 130; Indels 91; Gaps 13;

QY 58 ETVNNLNATLEOLKMAKTNLESAINQANTDKTTFDEHPNLVEAYKALKTKTLEQRTNL 117
DB 38 ELINHNNAISLSTEGSTDSINLGAQSPAVKSTRTTE-----LDVTCAAKTLTQTSVQK 93
QY 118 EGLSSTAYNQIRNLDVYNKASLTKTLDPLNGTGLDLSNEITANKNNINNTLTINE 177
DB 94 EMKVSQETQVSSE-----FSKRDVSNKEAVPVSKDELLEQSEVVYSTSIQKN-KILDN 148
QY 178 OKTNADALSFTKVKIQNNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKARTV- 236
DB 149 KKRANFTVTSPLIKKPSNKSQASGVINDS-----ASPLSKAKKEVVS 193
QY 237 -----WNGDEPSSRLTANSTIDVSIWYSLAGTNTKYQSFNSYGPS 279
DB 194 LRQPLKQKVEAQPILLISNSEKASVYTNSHDFWYQW-----DMKY---VTNNGES 243
QY 280 TGYLYFPYKLVKAADANNVGLQYKLNNGVQQVEFAFTSAN-----NT 323
DB 244 YA-LYQPSKKI-----SVGI---IDSGIMEHPDLSNLGNFYKLVKPGGFDNEPDE 293
QY 324 TANTTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIG-NIYLSNEN 379
DB 294 TGNPSDIVD-----KMGHGTEVAGQITANGNIGLVAPGITVNIYRVEGEN 338

RESULT 2
AR56_CANAL STANDARD; PRT; 857 AA.
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE ARG5,6 protein, mitochondrial precursor [Contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-L-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)].
GN ARG5,6.

```

```

OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 64385 / 1001;
RX MEDLINE-97195775; PubMed-9043106;
RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
RT Candida albicans.";
RL Microbiology 143:297-302(1997).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate -> N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate -> ADP + N-acetyl-L-
CC glutamate 5-phosphate.
CC -1- PATHWAY: SECOND AND THIRD STEPS IN ARGININE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: X98880; CAA67383.1; -
DR InterPro: IPR000706; AGPR_act_site.
DR InterPro: IPR001048; Akinase.
DR InterPro: IPR000534; Semialdh_dh.
DR Pfam: PF00696; akinase; 1.
DR Pfam: PF01118; Semialdehyde_dh; 1.
DR Pfam: PF02774; Semialdehyde_dhc; 1.
DR ProDom: PD003765; AGPR_act_site; 1.
DR PROSITE: PS01224; ARG; 1.
DR Oxidoreductase; transferase; Kinase; Arginine biosynthesis; NADP;
KW Mitochondrion; Multifunctional enzyme; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? ? ACETYLGLUTAMATE KINASE.
FT CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
FT ACT_SITE 659 669 BY SIMILARITY.
FT SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C471 CRC64;

Query Match 7.1%; Score 143; DB 1; Length 857;
Best Local Similarity 23.3%; Pred. No. 0.64;
Matches 99; Conservative 58; Mismatches 134; Indels 134; Gaps 25;

QY 46 IEASLSAY-----SRAETVNN---NLNATL-----EQLKMAKTNLESAINQANTD 88
DB 208 VEAAINSGLYPILTSLTASSTSGQLLVNADVAAGELAREFEPLKIVYLNKGGIINGNTG 267
QY 89 K-----TTFDNEHPNLV-EAYKALKTKTLEQ-----ATNLEGLSSTAYNQIRNLDVLYNK 138
DB 268 EKVSAINLDEEYEDLLKESWVYKTKLKEIHDLQHLPRSSVA-----IIDVNDL 320
QY 139 ASSLITKTLPLNGTGLDLSNEITANKNNIN-NTLTSTINEQKTNADALSFTKVKIQN 197
DB 321 QKELFTDS-----GAGTL-----IRRGYRLNRLNSLRDFG---NPDLRLNALLR----- 361
QY 198 EQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKARTVWNGDEPSSRIANT----- 250
DB 362 -----DPEIKTKGVSVASYLKFLDSVQFKS-----YGDPE-LEVLAIVVEQNDKI 405
QY 251 -----NSLTD-----VSWIYSLAGTNTKYQSFNSYNG-PSTGYL 283
DB 406 PKLDFLSKGTGLNNVTIDNIFNAIKKDYSLQVWVNDANLFWYFSKDSGSAKNGQI 465

```

```

Query Match      6.8%; Score 137; DB 1; Length 719;
Best Local Similarity 19.8%; Pred. No. 1.1;
Matches 105; Conservative 72; Mismatches 175; Indels 178; Gaps 23;

Qy 3 MSITKDKANPNNGQTQLBAARMEITDL-----INAKMTLAS----- 39
Db 38 MPTTLKDCYRNNSOVNEDAIIGNL-DLSLPIHTSPPTGESAGSGNATLRNDGNALDGG 96

```





Query Match									
Best Local Similarity 22.8%; Pred. No. 4;									
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;									
QY	7	KKDANPNNGQTOLFA-ARMELTDLINA-KAMTLASLDQYAKIEASLSAYSE-----AET	59						
DB	230	RKQAQADKKEDAEVKVREELGKLFSSKTAKGLDQEIQEHVKKETSSEENTQKVDEHYANS	289						
QY	60	VNNLNATLEOLKMAKNLNSAISAINQANTDKTTFNEHPNL-----VEAYKALKT	108						
DB	290	LONLAQKSELEEDKATYN-----EQATQVKQFLENQAQKLEIQIOLIKETNVKLYKAMSE	344						
QY	109	TLEQRATNLEGLSSTAYNQIRNNLVLYNKASSLITKTLPDPLNGTLLDSNEITANKNI	168						
DB	345	SLEQVEKELKHNSA-----NLEDLVAKSEIVREYEGKLQSKNLP--ELKOLEEEA	395						
QY	169	NNTLSTINEQ-----KTNADALSNSFIKKVI---QNNQSFVGTFTNANVOPSNYSFVAF	220						
DB	396	HSKLQGVVEDFRKKFKTSEQVTPPKRVARDLAANNENNOQKI-----ELTVSPENI	445						
QY	221	SADVTVPVYKYARRTVWNGDEPPSSRILANTNSITDVSIIYSLAGTNTKYQFSFS	275						
DB	446	-----IVYEGEDVKFTVTAKSQSKTTLDFSDLL-----TKYNPSVSDRISTN	487						
QY	276	YGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGNVQVQVEFATSTANNTAND-TPAV	331						
DB	488	YKTDNHNKIAEITIKNLKLNESQTVTLKAKDDSGNVVEKFTTIVQKKEQKVPTPEQ	547						
QY	332	DEIKVAKIV	340						
DB	548	KDSKTEERV	556						
RESULT 8									
RBP2_PLAVB STANDARD; PRT; 1251 AA.									
AC	Q00799;								
DT	01-APR-1993 (Rel. 25, Created)								
DT	01-APR-1993 (Rel. 25, Last sequence update)								
DT	01-OCT-1996 (Rel. 34, Last annotation update)								
DE	Reticulocyte binding protein 2 (Fragment).								
GN	RBP2.								
OS	Plasmodium vivax (strain Belem).								
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.								
OX	NCBI_TaxID=31273;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=92315338; PubMed=1617731;								
RA	Gallinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;								
RT	"A reticulocyte-binding protein complex of Plasmodium vivax								
RT	merozoites."								
RL	Cell 69:1213-1226(1992).								
CC	-1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO								
CC	HUMAN RETICULOCYTE CELLS.								
CC	-1- SUBCELLULAR LOCATION: Membrane-bound (Probable).								
CC									
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation								
CC	at the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>								
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).								
CC									
DR	EMBL; M88096; AAA29744.1;								
KW	Malaria; Receptor; Membrane.								
FT	NON_TER 1251 1251								
FT	SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;								
Query Match									
Best Local Similarity 6.4%; Score 129; DB 1; Length 1251;									
Best Local Similarity 21.1%; Pred. No. 6;									

Matches 89; Conservative 70; Mismatches 140; Indels 122; Gaps 20;									
QY	16	QTOLEARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAE-----TVNNLNAT	67						
DB	424	ETNLESYKHNLSITNIK-----QGGEKI-----YSKADIMQIKATSENTAETK	469						
QY	68	LEOLKMAKNLNSAISAINQANTDKTTFNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNQ	127						
DB	470	LEKVKDDQSNVYVNLQITIER-----NLIVTERNRLNGIDSTITNIEG---ALKE	517						
QY	128	IRNNL-----VDLYNKASSLITKTLPDPLNGTLLDSNEIT--TANKN	167						
DB	518	SKGNYEIGFLEKEEIGKRRKLVKITKS---INSTVG--NFSSLFNFDNLQYDFNKN	572						
QY	168	INNTLSITNEQKTNADALSNSF---IKVQNNNEQSFVGTFTNANVOPSNYSFVAFSADY	224						
DB	573	INDYENKGE-----IYNEFEGLSKISEN-----LRNASENTSDYN---SAKT	613						
QY	225	TPVNYKYARRTVWNGDEPPSSRILANTNSITDVSIIY-----SLAGTN---TKYQFSFSNYG	277						
DB	614	LRLEAQREKVNLLNKEEANKYLRDVKVKSFRIFNNKESLDKINEMIKKEQLTVNEGH	673						
QY	278	PSTGYLYPPYKLVKAADANNVGLQYKLNNGNVQVQVEFATSTANN-----TTA--	325						
DB	674	GNVQQLVENIK--ELVDENNLSDLKQATGKNEEQKITHSTLKNKAKTILGHVDTFSKY	731						
QY	326	-----NPTPAYDE-----IKVAK-----IVLSGLRFQONTIELSVPTGEGNMNV	365						
DB	732	VGKITPELALTELLGDAKLTAQELKEFSKNVNVLETNENSKNTNELDVHKNIQDAYKV	791						
QY	366	A	366						
DB	792	A	792						
RESULT 9									
WAPA_BACSU STANDARD; PRT; 2334 AA.									
AC	Q07833;								
DT	01-OCT-1994 (Rel. 30, Created)								
DT	01-OCT-1994 (Rel. 30, Last sequence update)								
DT	01-MAR-2002 (Rel. 41, Last annotation update)								
DE	Wall-associated protein precursor.								
GN	WAPA OR NI7G.								
OS	Bacillus subtilis.								
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;								
OC	Bacillus/Staphylococcus group; Bacillus.								
OX	NCBI_TaxID=1423;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=93302506; PubMed=8316082;								
RA	Poster S.J.;								
RT	"Molecular analysis of three major wall-associated proteins of								
RT	Bacillus subtilis 168: evidence for processing of the product of a								
RT	gene encoding a 258 kDa precursor two-domain ligand-binding								
RT	protein."								
RL	Mol. Microbiol. 8:299-310(1993).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RX	STRAIN=168 / BGSC1A1;								
RX	MEDLINE=95219088; PubMed=7704263;								
RA	Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;								
RT	"Cloning and sequencing of a 29 kb region of the Bacillus subtilis								
RT	genome containing the hut and wpa loci."								
RL	Microbiology 141:337-343(1995).								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RX	STRAIN=168 / BGSC1A1;								
RX	MEDLINE=97124196; PubMed=8969509;								
RA	Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.;								
RA	Miwa Y., Fujita Y.;								
RT	"Sequencing of a 65 kb region of the Bacillus subtilis genome								

containing the lic and cel loci, and creation of a 177 kb contig covering the gnt-sacxy region.";  
RL Microbiology 142:3113-3123(1996).  
CC -1- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,  
CC MOTILITY, SECRETION OR DIFFERENTIATION.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED  
CC INTO THE MEDIUM.  
CC -1- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE  
CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE  
CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED  
CC MOTIF REPEATED 31 TIMES.  
CC -1- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME  
CC SIMILARITY TO THE REPEAT IN E.COLI RHS GROUP OF PROTEINS (RHS-A-D).  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L05634; AAA22883.1; -;  
DR EMBL; D31856; BAA06556.1; -;  
DR EMBL; D29985; BAA06260.1; -;  
DR EMBL; D83026; BAA11683.1; -;  
DR EMBL; Z99124; CAB15959.1; -;  
DR PIR; S32920; S32920.  
DR Subtilist; BGL0797; wapa.  
DR InterPro; IPR003305; CBD\_6.  
DR Pfam; PF02018; CBD\_6; 1.  
DR Cell wall; Repeat; Signal; Complete proteome.  
FT SIGNAL 1 28  
FT CHAIN 29 2334  
FT DOMAIN 504 869  
FT REPEAT 504 605  
FT REPEAT 636 736  
FT REPEAT 769 869  
FT DOMAIN 1021 2139  
FT REPEAT 1021 1040  
FT REPEAT 1042 1061  
FT REPEAT 1063 1082  
FT REPEAT 1083 1102  
FT REPEAT 1109 1128  
FT REPEAT 1129 1148  
FT REPEAT 1150 1169  
FT REPEAT 1174 1193  
FT REPEAT 1199 1218  
FT REPEAT 1219 1238  
FT REPEAT 1646 1665  
FT REPEAT 1667 1686  
FT REPEAT 1690 1709  
FT REPEAT 1711 1730  
FT REPEAT 1732 1751  
FT REPEAT 1753 1772  
FT REPEAT 1795 1814  
FT REPEAT 1820 1839  
FT REPEAT 1840 1859  
FT REPEAT 1861 1880  
FT REPEAT 1887 1906  
FT REPEAT 1908 1927  
FT REPEAT 1929 1948  
FT REPEAT 1969 1982  
FT REPEAT 1983 2002  
FT REPEAT 2008 2027  
FT REPEAT 2028 2047  
FT REPEAT 2051 2070  
FT REPEAT 2071 2090  
FT REPEAT 2093 2112  
FT REPEAT 2120 2139  
SQ SEQUENCE 2334 AA; 258329 MW; B75138CCD278BAA3 CRC64;  
OR 32 (POTENTIAL).  
3 X 101 AA APPROXIMATE TANDEM REPEATS.  
1-1.  
1-2.  
1-3.  
31 X 21 AA APPROXIMATE TANDEM REPEATS OF  
X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).  
2-1.  
2-2.  
2-3.  
2-4.  
2-5.  
2-6.  
2-7.  
2-8.  
2-9.  
2-10.  
2-11.  
2-12.  
2-13.  
2-14.  
2-15.  
2-16.  
2-17.  
2-18.  
2-19.  
2-20.  
2-21.  
2-22.  
2-23.  
2-24. (APPROXIMATE).  
2-25.  
2-26.  
2-27.  
2-28.  
2-29.  
2-30.  
2-31.

Query Match 6.4%; Score 128.5; DB 1; Length 2334;  
Best Local Similarity 22.1%; Pred. No. 14;  
Matches 91, Conservative 56; Mismatches 152; Indels 113; Gaps 22;  
Qy 4 SITKDDANPNNGQTOLEARM-ELTDLINAKAMTULASLODYAKIEASLSAYSEATVN- 61  
Db 1772 SVANKEQNTTKRTFDKNRLTELDRGSGQTWTYPSDSKLLK---TFSWIHGDKQGTNQ 1828  
Qy 62 ---NNLNATLEQLMAKT-----NLESAINQANTDKTDEHNPV----- 100  
Db 1829 FTYNKLOMIEMKOSTSYSDYDENGWVQTFIT--GNGGTSFSYDERNLVSSLHIGDN 1887  
Qy 101 -----EAYK----ALKTTLEQRATLEGLSSTAYNQIRNNLVLYNKASSLITTLDP 150  
Db 1898 GGDILTESYEVADNGNRRTINSSAS---GKVOYEGKL-NQLVKETHEDGTVEITYDGF 1943  
Qy 151 NG----GFLD-SNEITANKNNNTLTSTINEQKTNADALSNSFIKKVIONNEQSFVGT 205  
Db 1944 GNRKTVTTIKDGSSTKVNASENIMNQLTKVANDESISYD-----KNGNRRTSDGKF 1992  
Qy 206 TNANVQPSNYSFVAFSADVTP-VNYKYARRTVWNGDEPSSRLANTNSITDVSWLYSLAG 264  
Db 1993 TYTWDADNLTAFTKKGDKFPATKY-----DEGNRIQKTVNGKV-TNIFYDGD 2043  
Qy 265 TTKYQFSFN-----YGPS-----TCYLFYFPYKLVKAADANNVGLQYKLN-NGN 308  
Db 2044 LNVLYETDADNNVTKSYTYGDSGQLLSYENGKKYF-----YHNAHGD 2087  
Qy 309 VOQVEFATSTSANNTA-----NP--TPAVDEIKVAKIVLSGLRFGONT 350  
Db 2088 I-----IAISDSTGKTVAKYQYDAMGNPTKTEASDEVKONRYAGYQYDEET 2135  
RESULT 10  
LMA\_DROME  
ID LMA\_DROME STANDARD; PRT; 3712 AA.  
AC Q00174;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Laminin alpha chain precursor.  
GN LANA OR LAMA.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93049203; PubMed=1425586;  
RA Kusche-Gullberg M., Garrison K., Mackrell A.J., Fessler L.I.,  
RA Fessler J.H.;  
RA "Laminin A chain: expression during Drosophila development and  
RA genomic sequence.";  
RL EMBO J. 11:4519-4527(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=94038678; PubMed=8223265;  
RA Henschel C., Garcia-Alonso L., Tang J., Goodman C.S.;  
RA "Genetic analysis of laminin A reveals diverse functions during  
RA morphogenesis in Drosophila.";  
RL Development 118:325-337(1993).  
RN [3]  
RP SEQUENCE OF 1762-3712 FROM N.A.  
RX MEDLINE=92078147; PubMed=1744083;  
RA Garrison K., Mackrell A.J., Fessler J.H.;  
RA "Drosophila laminin A chain sequence, interspecies comparison, and  
RA domain structure of a major carboxyl portion.";  
RL J. Biol. Chem. 266:22899-22904(1991).  
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF

CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
 CC -|- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.  
 CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC  
 CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE  
 CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES  
 CC IN CELL FATE AND PATTERN, MISSHAPEN LEGS AND DEFECTS IN WING  
 CC STRUCTURE.  
 CC -|- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
 CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT  
 CC MEMBRANES (MAJOR COMPONENT).  
 CC -|- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY  
 CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.  
 CC -|- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO  
 CC DEVELOPMENT AT 10-12 HOURS.  
 CC -|- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -|- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.  
 CC -|- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
 CC -|- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.  
 CC -|- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV) IS NOT  
 CC SIMILAR TO LAMININ DOMAIN IV).  
 CC -|- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M96388; AAA28662.1; -;  
 DR EMBL; L07288; AAC37178.1; -;  
 DR EMBL; M75882; AAA28661.1; -;  
 DR HSSP; P02468; ITLE.  
 DR FlyBase; FBgn0002526; LANA.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001886; LamNT.  
 DR InterPro; IPR000034; Laminin\_B.  
 DR InterPro; IPR002049; Laminin\_EGF.  
 DR InterPro; IPR001791; Laminin\_G.  
 DR InterPro; IPR001230; Prenyltn.  
 DR Pfam; PF00052; laminin\_B; 1.  
 DR Pfam; PF00053; laminin\_EGF; 20.  
 DR Pfam; PF00054; laminin\_G; 5.  
 DR Pfam; PF00055; laminin\_Nterm; 1.  
 DR PRINTS; PR00011; EGFLAMIN.  
 DR ProDom; PD002082; LamNT; 1.  
 DR ProDom; PD003031; Laminin\_B; 1.  
 DR SMART; SM00180; EGF\_Lam; 17.  
 DR SMART; SM00001; EGF\_Like; 1.  
 DR SMART; SM00281; LamB; 1.  
 DR SMART; SM00282; LamG; 5.  
 DR SMART; SM00136; LamNT; 1.  
 DR PROSITE; PS00022; EGF\_1; 17.  
 DR PROSITE; PS01186; EGF\_2; 5.  
 DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 19.  
 DR PROSITE; PS50025; LAM\_G\_DOMAIN; 5.  
 DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 3712 LAMININ ALPHA CHAIN.  
 FT DOMAIN 25 272 LAMININ N-TERMINAL (DOMAIN VI).  
 FT DOMAIN 273 332 LAMININ EGF-LIKE 1.  
 FT DOMAIN 333 402 LAMININ EGF-LIKE 2.  
 FT DOMAIN 403 447 LAMININ EGF-LIKE 3.  
 FT DOMAIN 448 494 LAMININ EGF-LIKE 4.  
 FT DOMAIN 495 540 LAMININ EGF-LIKE 5.  
 FT DOMAIN 541 586 LAMININ EGF-LIKE 6.

DOMAIN 587 631 LAMININ EGF-LIKE 7.  
 DOMAIN 632 676 LAMININ EGF-LIKE 8.  
 DOMAIN 677 731 LAMININ EGF-LIKE 9.  
 DOMAIN 732 784 LAMININ EGF-LIKE 10.  
 DOMAIN 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).  
 DOMAIN 816 1374 DOMAIN IV'.  
 DOMAIN 1375 1420 LAMININ EGF-LIKE 12.  
 DOMAIN 1421 1465 LAMININ EGF-LIKE 13.  
 DOMAIN 1466 1513 LAMININ EGF-LIKE 14.  
 DOMAIN 1514 1564 LAMININ EGF-LIKE 15.  
 DOMAIN 1565 1574 LAMININ EGF-LIKE 16 (N-TERMINAL).  
 DOMAIN 1575 1775 LAMININ DOMAIN IV (DOMAIN IV).  
 DOMAIN 1776 1808 LAMININ EGF-LIKE 16 (C-TERMINAL).  
 DOMAIN 1809 1858 LAMININ EGF-LIKE 17.  
 DOMAIN 1859 1916 LAMININ EGF-LIKE 18.  
 DOMAIN 1917 1969 LAMININ EGF-LIKE 19.  
 DOMAIN 1970 2016 LAMININ EGF-LIKE 20.  
 DOMAIN 2017 2063 LAMININ EGF-LIKE 21.  
 DOMAIN 2064 2111 LAMININ EGF-LIKE 22.  
 DOMAIN 2112 2671 DOMAIN II AND I.  
 DOMAIN 2672 2868 LAMININ G-LIKE 1.  
 DOMAIN 2876 3048 LAMININ G-LIKE 2.  
 DOMAIN 3055 3223 LAMININ G-LIKE 3.  
 DOMAIN 3249 3528 LAMININ G-LIKE 4.  
 DOMAIN 3534 3709 LAMININ G-LIKE 5.  
 DOMAIN 2178 2249 COILED COIL (POTENTIAL).  
 DOMAIN 2301 2321 COILED COIL (POTENTIAL).  
 DOMAIN 2376 2450 COILED COIL (POTENTIAL).  
 DOMAIN 2541 2676 COILED COIL (POTENTIAL).  
 DOMAIN 3270 3296 POLY-THR.  
 DISULFID 273 282 BY SIMILARITY.  
 DISULFID 275 296 BY SIMILARITY.  
 DISULFID 298 307 BY SIMILARITY.  
 DISULFID 310 330 BY SIMILARITY.  
 DISULFID 333 342 BY SIMILARITY.  
 DISULFID 335 367 BY SIMILARITY.  
 DISULFID 370 379 BY SIMILARITY.  
 DISULFID 382 400 BY SIMILARITY.  
 DISULFID 403 414 BY SIMILARITY.  
 DISULFID 405 421 BY SIMILARITY.  
 DISULFID 423 432 BY SIMILARITY.  
 DISULFID 435 445 BY SIMILARITY.  
 DISULFID 448 460 BY SIMILARITY.  
 DISULFID 450 468 BY SIMILARITY.  
 DISULFID 470 479 BY SIMILARITY.  
 DISULFID 482 492 BY SIMILARITY.  
 DISULFID 495 507 BY SIMILARITY.  
 DISULFID 497 514 BY SIMILARITY.  
 DISULFID 516 525 BY SIMILARITY.  
 DISULFID 528 538 BY SIMILARITY.  
 DISULFID 541 553 BY SIMILARITY.  
 DISULFID 543 560 BY SIMILARITY.  
 DISULFID 562 571 BY SIMILARITY.  
 DISULFID 574 584 BY SIMILARITY.  
 DISULFID 587 599 BY SIMILARITY.  
 DISULFID 589 605 BY SIMILARITY.  
 DISULFID 607 616 BY SIMILARITY.  
 DISULFID 619 629 BY SIMILARITY.  
 DISULFID 632 644 BY SIMILARITY.  
 DISULFID 634 650 BY SIMILARITY.  
 DISULFID 652 661 BY SIMILARITY.  
 DISULFID 664 674 BY SIMILARITY.  
 DISULFID 677 691 BY SIMILARITY.  
 DISULFID 679 700 BY SIMILARITY.  
 DISULFID 702 711 BY SIMILARITY.  
 DISULFID 714 729 BY SIMILARITY.  
 DISULFID 732 746 BY SIMILARITY.  
 DISULFID 734 753 BY SIMILARITY.  
 DISULFID 755 764 BY SIMILARITY.  
 DISULFID 767 782 BY SIMILARITY.  
 DISULFID 1375 1387 BY SIMILARITY.  
 DISULFID 1377 1394 BY SIMILARITY.  
 DISULFID 1396 1405 BY SIMILARITY.



NCBI\_TaxID=4932;  
[1]  
QY SEQUENCE FROM N.A.  
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,  
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;  
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
[2]  
QY SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,  
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
[3]  
QY FIBRONECTIN TYPE III DOMAIN.  
RA Bateman A., Chothia C.;  
RX MEDLINE=97148176; PubMed=8994808;  
RA "Fibronectin type III domains in yeast detected by a hidden Markov  
model";  
RL Curr. Biol. 6:1544-1546 (1996).  
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U18779; AAB64999.1;  
DR PIR; S30834; S30834.  
DR SGD; S0000769; YEL043W.  
DR InterPro; IPR003961; FN\_III.  
DR Pfam; PF00041; fn3; 1.  
DR SMART; SM00060; FN3; 1.  
KW Hypothetical protein.  
FT DOMAIN 35 125 FIBRONECTIN TYPE-III.  
SQ SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;  
  
Query Match 6.3%; Score 127; DB 1; Length 956;  
Best Local Similarity 19.7%; Pred. No. 5.5;  
Matches 104; Conservative 86; Mismatches 163; Indels 176; Gaps 25;  
QY 5 ITRKDANPNQTOLEAARMELTDL-----INA-RAMTLASLDQYAKIEASLSAYSEA 57  
DB 337 LNESASVANINKEIESLQNEISKEESRNLNASKSLITSIVNANVENDKPIASGEL 396  
QY 58 ETYNNLN-ATLE-----QLKMAKTLESAINANTDKTTFDNEHNLVEAYKALK 107  
DB 397 SAVLKUNLDFLENGFSLNAGEEFLSKLADSLIKWIKQELSIOE---LEANWKLQR 453  
QY 108 TLEQRATNLEGLSSATYNOIRNVLNLYNKASSLIPTKL-DPL--NGGILLDSNETT 164  
DB 454 SNLLKKISALE-----NQF--NEMSLNLR--NLKTKLMVQPYKKNNGDSLAATNSNSA 502  
QY 165 NKN-----INNTLSTINEQKTNADAL----- 185  
DB 503 EKNRSSGSIQLPLSNMRSRTGSIDLNNKNSIN--NSNADAPPLRLHNPVSPSNEP 560  
QY 186 ---SNSETKKIQNNE--QSFVGTETNAN-----VQPSNY SFVA-----FSADYTPVNYK 230  
DB 561 TQPSSLSQLTQTDIDNLSMLNHLSSNNENKQPPSSVSHALPTATANATATATATN-G 619  
QY 231 YARRTVWNG---DEPSSKILA-----NTNSITDVSHIYSLAGTNTKYQSFNSYNG 277  
DB 620 HSRSLNLTWTAQAPQSHQOVSTELDQAFEDYDANHL--ISGLQNMIDYDTPDNISNYS 677

QY 278 -----PSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQ 310  
DB 678 KGFTTDELNDYTKQOPQVRSSTNESLFTGTPMSSYK-----ANPVISPY--SSHLR 729  
QY 311 QVEFATSTANNTTAMPTPAVDIKVAKIVLSGLRF-----GQNTIELS-- 354  
DB 730 QTSNATNTNPMHPQSLAATLNDPSLQSFVRSGFSYSAFPANSLQNNINGNETENISPR 789  
QY 355 -----VPTGEGNNKVPAMIGNIYLSSNENNADKI 384  
DB 790 ISSDFNLLVNLSPRLSNDVPIVPGNNTLTTPSHSNLLTMNHQPTADNI 838  
  
RESULT 13  
PSTL\_YEAST STANDARD; PRT; 444 AA.  
ID PSTL\_YEAST  
AC Q12355;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Protoplast secreted protein 1 precursor.  
GN PST1 OR YDR055W OR D4214 OR YD9609.09  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96381250; PubMed=8789263;  
RA Brandt P., Ramlow S., Otto B., Bloeker H.;  
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm  
RT of Saccharomyces cerevisiae chromosome IV";  
RL Yeast 12:85-90(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP IDENTIFICATION.  
RX MEDLINE=99251092; PubMed=10234784;  
RA Pardo M., Monteoliva L., Pla J., Sanchez M., Gil C., Nombela C.;  
RT "Two-dimensional analysis of proteins secreted by Saccharomyces  
RT cerevisiae regenerating protoplasts: a novel approach to study the  
RT cell wall";  
RL Yeast 15:459-472(1999).  
RN [4]  
RP GPI-ANCHOR.  
RX MEDLINE=20469049; PubMed=11016834;  
RA Terashima H., Yabuki N., Arisawa K., Hamada K., Kitada K.;  
RT "Up-regulation of genes encoding glycosylphosphatidylinositol  
RT (GPI)-attached proteins in response to cell wall damage caused by  
RT disruption of FKS1 in Saccharomyces cerevisiae";  
RL Mol. Gen. Genet. 264:64-74(2000).  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND  
CC SECRETED BY REGENERATING PROTOPLASTS.  
CC -!- SIMILARITY: BELONGS TO THE SPS2 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X84162; CAA58971.1;  
DR EMBL; Z74351; CAA98873.1;  
DR EMBL; Z49209; CAA98084.1;  
DR SGD; S0002462; PST1.  
DR COMPLEYEST-2DPAGE; Q12355;  
KW Glycoprotein; Membrane; GPI-anchor; Signal.  
FT SIGNAL 1 19  
POTENTIAL.

FT	CHAIN	20	?	PROTOPLAST SECRETED PROTEIN 1.		
FT	PROPEP	?	444	REMOVED IN MATURE FORM (POTENTIAL).		
FT	DOMAIN	356	416	SER-RICH.		
FT	CARBOHYD	57	57	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	76	76	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	83	83	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	86	86	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	210	210	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	228	228	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	235	235	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	242	242	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	263	263	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	268	268	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	280	280	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	305	305	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD	329	329	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	SEQUENCE	444 AA;	45776 MW;	230F60CACA5921A4 CRC64;		
Query Match 6.3%; Score 126.5; DB 1; Length 444;						
Best Local Similarity 18.1%; Pred. No. 2.3;						
Matches 73; Conservative 79; Mismatches 146; Indels 105; Gaps 15;						
QY	17	TOLEAARWE-LTDLINAKAMT	LASLDYAKTEASLS	SAYSEAEVTNNNLNATL	EQLKMAK 75	
DB	88	TSFAADLESITDSLN	IQSITLT	-----SASFGSLQVDSIK	LTPAISFT 136	
QY	76	TNLESAINQANTOKT	TFDNEHPNLVEAY	KALKTTLEQRATN	LEGSS--TAYNOIRNLV 133	
DB	137	SNIKSANIYISDTSL	-----QSVDFGSALK	KVNVFVNNNKKLT	SIKSPVETVDSLQ 190	
QY	134	DLYNKASSLITK	LDPL---NGGTLDSNEIT	TAKN-----INNTLS	INEQKINA 182	
DB	191	FSFENGNTKI-	TFDDLWANNISL	TDVHSVFANQLK	INSSSLFNINSLFTKLNT 248	
QY	183	-----DALSNSFIK	-----VTQNN-----	-----BQSPVGF	FTN 207	
DB	249	IGQTFSTVSNDY	LKNLSFNLSL	TIGGALVANNFTGLQ	IGGLDNLTTIGGTVLVVGNFTS 308	
QY	208	ANYQPSNYSVAF	SADYTPVNYKYARR	TWNGDEPSSRILANT	NSITDVSWIYSLAGTNT 267	
DB	309	LNL--DSLKSVKG	ADVESKSSNFSCAL	-----KALQKGIG	KEGSEVCKNGASST 358	
QY	268	KYQFSFSGNY	PGSTGLYFPYK	KAADANNVGLYK	LNNGNVQVQVEFATST	SANNNTANP 327
DB	359	SVKLSSTSK	QSQSS-----QTTAKV	SKSSSKAEKKFTS	GDGIKAASASSVSSGASS-- 410	
QY	328	TPAVDEIKVAK	IVLGLURFGONT	IELSVPTGEGNNK	VAPMIG 370	
DB	411	-----SSSKSKG	NAAIWAPIGO	-----TTPIVG	434	

RESULT	14	
MSPI_PLAFK		
ID	MSPI_PLAFK	STANDARD; PRT: 1630 AA.
AC	P04932;	
DT	13-AUG-1987 (Rel. 05, Created)	
DT	01-FEB-1996 (Rel. 33, Last sequence update)	
DT	01-OCT-1996 (Rel. 34, Last annotation update)	
DE	Merzolate surface protein 1 precursor (Merzolate surface antigens) (PMMSA) (P190).	
GN	MSP-1.	
OS	Plasmodium falciparum (isolate K1 / Thailand).	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX	NCBI_TaxID:5839;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=86135024; PubMed=3004972;	
RA	MacKay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,	
RT	Stunnenberg H., Bujaard H.;	
RT	*polymorphism of the precursor for the major surface antigens of	

Plasmodium falciparum merozoites: studies at the genetic level.";  
EMBO J. 4:3823-3829(1985).  
[2]  
RL REVISIONS, SEQUENCE FROM N.A.  
RN  
RP Pen W., Tolle R., Bujard H.;  
RA Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RL  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
CC (- potential).  
CC  
CC -1- PDM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42  
CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF  
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X03371; CAA27070.1; -.  
DR  
DR PIR; A25120; SAZQK1.  
DR  
DR InterPro: IPR000561; EGF-like.  
DR  
DR Pfam; PF00008; EGF; 1.  
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor.  
KW  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.  
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.  
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.  
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT SEQUENCE 1630 AA; 187289 MW; ABBDEC3CE0A46322 CRC64;  
SQ

Query Match 6.3%; Score 126.5; DB 1; Length 1630;  
Best Local Similarity 20.7%; Pred. No. 11;  
Matches 86; Conservative 62; Mismatches 155; Indels 113; Gaps 19;

Qy	37	LASLDYAK-----IRASUSSAYSEAEVTNNUNLNATLEQLKMAKTNWLSAINQANTDKTT	91
Db	226	VGMEDYIKKNKKTNIENIELIESKKTIDKNKNATKEEK-----KKLYQAQYDLST	278
Qy	92	FDNEHPNLVPEAYKALKTTLLEORATNLEGUSSTAYNOIRNNLVLYNKASSLTKTLDPLN	151
Db	279	YNKQ---LEEAHN-LISVLEKRTDTLKK-----NENIKELDKINEI--KNPPAN	323
Qy	152	GG---TLLDSN-EITTANKNINNTLSTINEOKTNDAL-----SNSFTK	191
Db	324	SGNTPNLLDKNKKIEBEHEKEIEIAKTI---KFNIDSUFTDPLEYLYREKNNKIDS	380
Qy	192	KVQNNQESVPGTTFNANVOPSNYSFV-----AFSADVTPVNYKYARRTWNGDEPS	243
Db	381	AKVETKESTEPNEYPNGVTYPLSYDINNALNELNSFGDLPINPFDY-----KEPS	431
Qy	244	SRILAN-----TNSITDVSNIYSLAGNTKYQFS	272
Db	432	KNIVTDNERKKFNEIKETIKIEKKIESDKSYEDRSKSLNDITKEYEKL-LNEYIDSK	490
Qy	273	FSNYGSGTGY-----LYFPYKLVKADANNVGLQYKLNINGVQOVFEATSTSANNTTAMP	327
Db	491	FNNNIDITNPEKMMGRKYSYQVEKLKTH--HNTFASYPENSKHNLEKLTALKYMEDYSLRN-	548

Search completed: June 12, 2002, 10:51:09  
Job time: 207 sec

Query Match	6.3%;	Score 126.5;	DB 1;	Length 1639;
Best Local Similarity	20.7%;	Pred. No. 11;		
Mismatches	62;	Mismatches 155;		
Conservative	96;	Indels 113;		Gaps 19;



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:50:40 ; Search time 48.58 Seconds  
(without alignments)  
1403.047 Million cell updates/sec

Title: US-09-147-052-4\_COPY\_693\_1086

Perfect score: 2004

Sequence: 1 GCMSTTKDANPNNGQTGLE.....SSNNADKIPGYRRDGTGL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPREMBL19.\*
- 2: sp\_archaea.\*
- 3: sp\_bacteria.\*
- 4: sp\_fungi.\*
- 5: sp\_human.\*
- 6: sp\_invertebrate.\*
- 7: sp\_mhc.\*
- 8: sp\_mammal.\*
- 9: sp\_organelle.\*
- 10: sp\_phage.\*
- 11: sp\_plant.\*
- 12: sp\_rodent.\*
- 13: sp\_virus.\*
- 14: sp\_vertebrate.\*
- 15: sp\_unclassified.\*
- 16: sp\_rvirus.\*
- 17: sp\_bacteriap.\*
- 18: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1200	59.9	271	2 Q49464	Q49464 mycoplasma
2	837	41.8	671	2 Q918D6	Q918D6 mycoplasma
3	831.5	41.5	702	2 Q49499	Q49499 mycoplasma
4	826	41.2	632	2 Q9XCG8	Q9XCG8 mycoplasma
5	805	40.2	584	2 Q9KH13	Q9KH13 mycoplasma
6	761	38.0	680	2 Q9KH14	Q9KH14 mycoplasma
7	754	37.6	702	2 Q49498	Q49498 mycoplasma
8	746	37.2	645	2 Q9ZID1	Q9ZID1 mycoplasma
9	745	37.2	644	2 Q9ZHR9	Q9ZHR9 mycoplasma
10	744	37.1	650	2 Q49495	Q49495 mycoplasma
11	743	37.1	649	2 Q49497	Q49497 mycoplasma
12	713	35.6	647	2 Q49468	Q49468 mycoplasma
13	704	35.1	656	2 Q9KH15	Q9KH15 mycoplasma
14	690.5	34.5	703	2 Q05122	Q05122 mycoplasma
15	619.5	30.9	419	2 Q918D5	Q918D5 mycoplasma
16	575.5	28.7	386	2 Q49500	Q49500 mycoplasma

17	557	27.8	367	2 Q9XCG7	Q9XCG7 mycoplasma
18	189.5	9.5	320	2 Q49496	Q49496 mycoplasma
19	175	8.7	6713	16 Q99054	Q99054 staphylococ
20	175	8.7	6713	16 Q931R6	Q931R6 staphylococ
21	173	8.6	1302	2 Q49547	Q49547 mycoplasma
22	164.5	8.2	2481	16 Q99QR6	Q99QR6 staphylococ
23	163.5	8.0	3890	16 Q99U53	Q99U53 staphylococ
24	160.5	8.0	433	16 Q97TB6	Q97TB6 clostridium
25	160.5	8.0	682	2 Q48674	Q48674 lactococcus
26	159	7.9	933	2 Q53653	Q53653 staphylococ
27	158.5	7.9	1344	2 Q49545	Q49545 mycoplasma
28	156	7.8	807	5 Q96262	Q96262 plasmodium
29	155	7.7	1072	16 Q9CF64	Q9CF64 lactococcus
30	155	7.7	4688	16 Q9PQ08	Q9PQ08 ureaplasma
31	154.5	7.7	1365	2 Q49525	Q49525 mycoplasma
32	153.5	7.7	1795	2 Q9LJC9	Q9LJC9 staphylococ
33	153.5	7.7	2478	2 Q9RL69	Q9RL69 staphylococ
34	151.5	7.6	1237	16 Q9ZKS7	Q9ZKS7 helicobacte
35	151.5	7.6	2478	2 Q9LCH2	Q9LCH2 staphylococ
36	151	7.5	624	2 Q49548	Q49548 mycoplasma
37	150.5	7.5	604	5 Q26021	Q26021 plasmodium
38	150.5	7.5	1051	2 Q49524	Q49524 mycoplasma
39	149.5	7.5	962	2 Q49546	Q49546 mycoplasma
40	148	7.4	716	2 Q49526	Q49526 mycoplasma
41	147.5	7.4	655	5 Q26109	Q26109 plasmodium
42	146	7.3	654	2 Q9RLP9	Q9RLP9 mycoplasma
43	145	7.2	2399	16 Q9ZKS9	Q9ZKS9 helicobacte
44	144	7.2	1107	2 Q9F2D8	Q9F2D8 salmonella
45	143.5	7.2	751	3 Q94358	Q94358 schizosacch

ALIGNMENTS

RESULT 1

Q49464	PRELIMINARY;	PRT;	271 AA.
ID Q49464;			
AC Q49464;			
DT 01-NOV-1996 (TREMBLrel. 01, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE TM-1 (FRAGMENT).			
GN TM-1.			
OS Mycoplasma gallisepticum.			
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;			
OC Mycoplasmataceae; Mycoplasma.			
OX NCBI_TaxID=2096;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=94025893; PubMed=8212828;			
RA Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,			
RA Kanogawa K., Aoyama S., Iritani Y., Hayashi Y.;			
RT "Cloning and DNA sequence of a 29 kilodalton polypeptide gene of			
RL Mycoplasma gallisepticum as a possible protective antigen.";			
RL Vaccine 11:1061-1066(1993).			
DR EMBL: S65869; AAB28343.2; -			
FT NON_TER	271	271	
SQ SEQUENCE	271 AA;	29817 MW;	8B25DE0CD5C85CA2 CRC64;

Query Match 59.9%; Score 1200; DB 2; Length 271;  
Best Local Similarity 97.1%; Pred. No. 2.2e-51;  
Matches 238; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy	2	CMSITTKDANPNNGQTGLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN	61
Db	27	CMSITTKDANPNNGQTGLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN	86
Qy	62	NNLNRATLEQLKMAKTNLSEAINQANTDKTTFDNEHPNLVEAYKALKTLEQRATNLEGLS	121
Db	87	NNLNRATLEQLKMAKTNLSEAINQANTDKTTFDNEHPNLVEAYKALKTLEQRATNLEGLA	146
Qy	122	STAYNQIRNNLVLDLYNKASSLITKTLDPLNGCTLLDSNEITTANKNINNTLSTINEQKTN	181

|||||  
Db 147 STAYNOIRNNLDVLYNNASSLITKLDPLNGMLDSNEITVVRNINNLSTINEOKTN 206  
QY 182 ADALNSFIKKVIONNEQSGVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241  
Db 207 ADALNSFIKKVIONNEQSGVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 266  
QY 242 PSSRI 246  
Db 267 PSSRI 271  
RESULT 2  
Q9L8D6 PRELIMINARY; PRT; 671 AA.  
AC Q9L8D6;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE PMGA-LIKE PROTEIN 9.2.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F;  
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,  
RA May J.D., Hughlett M.B.;  
RT "A novel PMGA-like gene from the F-strain (vaccine strain) of  
RT Mycoplasma gallisepticum";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF210770; AAF29524.1; -;  
SQ SEQUENCE 671 AA; 71898 MW; 4D6AE6B59175D679 CRC64;  
Query Match 41.8%; Score 837; DB 2; Length 671;  
Best Local Similarity 43.4%; Pred. No. 2.5e-33;  
Matches 179; Conservative 75; Mismatches 118; Indels 40; Gaps 9;  
QY 9 DANP-----NNGQLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEAEVTVN 62  
Db 55 DTNPGNDGGMENSAQOOLAAAKKELSDLLATQNSLNKAKYADYTNQNTLTAAVTTAKSTD 114  
QY 63 NLNATLEOLKMAKTNLESAINOANTDKTTFDNEHNPVAYKALKTTLPORATNLEGLSS 122  
Db 115 NTSVLEQVKSATSTLQAADTAASKTSFDEKNPELIKAYALKETLKNEETVLSGLTD 174  
QY 123 TAYNOIRNNLDVLYNKASSLITKLDPLNGGTLDSNEITTANKNNINLTSTINEOKTNA 182  
Db 175 SNFATIKTNLTALYQSGKDFVKATLDPVSGNA-PQIADITKADKOIADAVSKLETWKTNA 233  
QY 183 DALNSFIKKVIONNEQSGVGTFTNANVQPSNYSFVAFSADVTP-----VNYKYARRTVN 238  
Db 234 NTLATSFVKVLYKNGKLTGIDT-TNNRQPGNYSFVGYSVNATNNNEINPNWPAQRKVT 292  
QY 239 GDBPSSRILA----NTNSTIDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAAD 294  
Db 293 SDNGRTLSISSDSDNSTITEVSWIYSLGAGTKYSLTFNYGYGPSTGYLYFPYKLVKEGD 352  
QY 295 ANNVGLQYKLNGNVOQVEFA-----TSTSA-----NNTT-----ANPTPAVD 332  
Db 353 ENNVALQYTLNNGSAQEVNFAFPVKTSVSDSGDSNNOTESAAATWPTVSTDLNPAPTWS 412  
QY 333 EIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNNENADKI 384  
Db 413 DINIAKLTLSNLKFGSNTIEFSVPTPEFS--NKVAPMIGNMYLTSTIANEAKV 462  
RESULT 3  
Q49499 PRELIMINARY; PRT; 702 AA.  
ID Q49499  
AC Q49499;

DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE PMGAI.4 PROTEIN PRECURSOR.  
GN PMGAI.4.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S6;  
RX MEDLINE=95010739; PubMed=7925999;  
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
RA Browning G.F., Whithear K.G., Walker I.D.;  
RT "The organisation of the multigene family which encodes the major cell  
RT surface protein, pmga, of Mycoplasma gallisepticum";  
RL FEBS Lett. 352:347-352(1994).  
DR EMBL; L28424; AAA62418.1; -;  
KW SIGNAL.  
FT SIGNAL 1 25 POTENTIAL.  
SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;  
Query Match 41.5%; Score 831.5; DB 2; Length 702;  
Best Local Similarity 42.0%; Pred. No. 4.8e-33;  
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;  
QY 9 DANPNG-----QPLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEAEVTVN 62  
Db 68 NTPNGNGGGTDNAOOLAAAKKELSDLLATQNSLSTYADYANIONTLTAATYTTAKSTD 127  
QY 63 NLNATLEOLKMAKTNLESAINOANTDKTTFDNEHNPVAYKALKTTLPORATNLEGLSS 122  
Db 128 NTSVLEQVKSATSTLQAADTAASKTSFDEKNPELIKAYALKETLKWRNLSGLTD 187  
QY 123 TAYNOIRNNLDVLYNKASSLITKLDPLNGGTLDSNEITTANKNNINLTSTINEOKTNA 182  
Db 188 SNFATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAVSQANTNISNAVSKLETWKTNA 246  
QY 183 DALNSFIKKVIONNEQSGVGTFTNANVQPSNYSFVAFSADVTP-----VNYKYARRTV 236  
Db 247 TVLATSFVKVLYKNGKLTGIDT-TNNRQPGNYSFVGYSDVTTGSDNAPNWSFAQRV 305  
QY 237 WNGD-----EPSRILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVK 291  
Db 306 WTSNTDILSQPOPAEGENQOQAPDVSWIYNLTGMAKYSITENYGYGPSTGYLYFPYKLVN 365  
QY 292 ADANNVGLQYKLNGNVOQVEFATS-----TSANN-----TTANPTPAVD 332  
Db 366 SSDSDKVALLEYKLNESAVKTIDFSPQTSFVADATRENNRSTAAAPAQGSTINPAPTLD 425  
QY 333 EIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNNENADK 383  
Db 426 DIKIAKLTLSNLKFGSNTIEFSVPTTAKETSKVAPMIGNMYLTSSDROVVK 477  
RESULT 4  
Q9XCG8 PRELIMINARY; PRT; 632 AA.  
AC Q9XCG8;  
DT 01-NOV-1999 (TReMBLrel. 12, Created)  
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE VLHAI PRECURSOR (FRAGMENT).  
GN VLHAI.  
OS Mycoplasma imitans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=29560;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=4229;

```
RX MEDLINE=99392472; PubMed=10463176;
RA Markham P.F., Duffy M.F., Glew M.D., Browning G.F.;
RT "A gene family in Mycoplasma imitans closely related to the pmgA
family of Mycoplasma gallisepticum.";
RL Microbiology 145:2095-2103(1999).
DR EMBL; AF141940; AAD39483.1; -.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 17 POTENTIAL.
FT CHAIN 18 632 VLHAI.
FT SEQUENCE 632 AA; 66959 MW; 173F5B12E705BE47 CRC64;

Query Match 41.2%; Score 826; DB 2; Length 632;
Best Local Similarity 45.3%; Pred. No. 8e-33;
Matches 183; Conservative 65; Mismatches 120; Indels 36; Gaps 9;

Qy 8 KDAPNNGQT-----OLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSAE 58
Db 34 KPNPDTGNTNPGDGTDAKQADAKASLNTLLGSGSTNVALYEDYAKIKDTLSSAYSAQ 93
Qy 59 TVNNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTATNLE 118
Db 94 TTANNANATLKDNDKAAALQTAISDAANAKTEFDKANGLVSAKAKUKETLKSETTNDL 153
Qy 119 GLSSTAYNQIRNNLVLYNKASSLITTKTLDPLNGGTLTDSNEITTANKNINNTLSTINEQ 178
Db 154 GLSQDNTSAIKANLVLYNKAKDETTVTLDLP-TSGMIPKVDIEITSANTAIQAYSAIDSQ 212
Qy 179 KTNADALSNTFKKVIQNNQSFVGTFTN-----ANVQSNYSFVAFSADVTP----- 226
Db 213 KTNADTAATTEIKELKLSAKLT-PCGTAEGQQAASQCPGNYSFVGSNDVTTGRTGSQEQ 271
Qy 227 --VNYKYARRVWNGDEPS---SRLATNNTSITDVSWIYSLAGTNTKYQSFSGYGPSTG 281
Db 272 DLPSNWFNAKRVWTEGLSAGQTLQVSSTPLTLDVSWIYSLTGAGSKYTLTFTYGPSTA 331
Qy 282 YLYFPYKLVKAADANNVGLQYKLNNG-NVQOVEFATTSANNTTANPTPAYDEIKVAKIV 340
Db 332 YLYFPYKLVQSSDSNKIGLQYKLNNSLTVPTFGNETN-----DSGATPAIDDIKVEKVT 387
Qy 341 LSGLRFQNTIELSVPTGEGNNKVPAMIGNIYSSNNENNADKI 384
Db 388 LSNLNFQDNTIEFSVATEE---NKVAPMIGNMYLTSTSTNNVDKI 428

RESULT 5
Q9KH13 PRELIMINARY; PRT; 584 AA.
AC Q9KH13;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ADHESIN PMGAL.4 (FRAGMENT).
GN PMGAL.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmgA multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91415.1; -.
DR InterPro; IPR001986; EPSP_synthase.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
FT NON_TER 584 584
FT SEQUENCE 584 AA; 62453 MW; 5C467BA55FB27A72 CRC64;
```

```
Query Match 40.2%; Score 805; DB 2; Length 584;
Best Local Similarity 42.4%; Pred. No. 7.6e-32;
Matches 180; Conservative 72; Mismatches 113; Indels 60; Gaps 10;

Qy 11 NPNGQT-----OLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVNNLNA 66
Db 69 NPNSGNTTPEQQLAAARKTLTDLTGENTNVALYADYAKIQSTLSLAYMTAKTASENTSA 128
Qy 67 TLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTATNLEGLSSTAYN 126
Db 129 TLDNLSASTTLQAADKAASNKRTFDSANQPLVTAYNQLKTTLOSKTTSLEGLSENKYS 188
Qy 127 QIRNNLVLYNKASSLITTKTLDPLNGGTLTDSNEITTANKNINNTLSTINEQKTNADALS 186
Db 189 SIKNHLSKLFDAGSAIAARTLDP-TMGIVPEVMSYTKANEDIMTAVSKLTEWKNADKF- 246
Qy 187 NSFIKKVIQNNQSFVGTFTNAN--VQPSNYSFVAFSADVTP-----VNYKYARTT 235
Db 247 NDFEKKPL--SKEKLVTNDRAHNOEQPANWFSAGYSVDLTGTGSGNSQNLPNWNFQARK 304
Qy 236 VYNGD---EPSSRIILANTNSITDVSWIYSLAGTNTKYQSFSGYGPSTGYLPYKLVKA 292
Db 305 VMTSEGOQTGKALVSSPVSATDVSWIYSLAGETKYLTSFEYGPDTAFLIFPKLVKQ 364
Qy 293 ADANNVGLQYKLNNGNVQOVEP-----ATST-----S 319
Db 365 ADSSVALQYSLNKTSSKLINEPEAKTPTNADQSENGVATTSTTEGRSSSEVLVADEYA 424
Qy 320 ANNTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYSSNNEN 379
Db 425 AVNEMNPTPTVDINIAKVTLSGLTGFGTNTIEFSVPT-----NKVAPMIGNMYLTSSNG 479
Qy 380 NADKI 384
Db 480 SQGKI 484

RESULT 6
Q9KH14 PRELIMINARY; PRT; 680 AA.
AC Q9KH14;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ADHESIN PMGAL.3.
GN PMGAL.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmgA multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91414.1; -.
SQ SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DFE47 CRC64;

Query Match 38.0%; Score 761; DB 2; Length 680;
Best Local Similarity 39.8%; Pred. No. 1.2e-29;
Matches 164; Conservative 71; Mismatches 129; Indels 48; Gaps 7;

Qy 9 DANPNNGQTQLEAARNE-----LTDLINAKAMTILASLDYAKIEASLSAYSEATVYN 61
Db 63 DTPNGGGQGMNATNQELVNAKKAALSDLLGGESKVELYADYAKIKADITLISAYAKTTS 122
Qy 62 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTATNLEGLS 121
Db 123 DSSTLTDQVKATATLTQTAINTAASDKKFDQSQSQLLMAYKVLKDLINKKEAIVMSLN 182
```

QY	122	STAYNOIRNNLDVLYNKASSLLITKTDPLNGGTLSDSNEITTANKNIINNTL-----ST	174
Dd	183	QEKYSAILSEINAASATAFEIVKLTLPVNG-----NLPPVAALNAENTKILEAIKEEK	236
QY	175	INEOKTNADALSNSFTKKVIQNNEQSFVCGETNNAVNPQSNTSVFAFSADVTPPNVKYARR	234
Dd	237	INSEKSNADLFANYQLYLK---DRTKLMEGSNNTKPGNTSFVAYADIASPWNFAQR	293
QY	235	TVMWGD-----BPSSRILANTNSITDVSWIYSLAGTNTKYQFSFSGYGSTGLYFPYKL	289
Dd	294	TWTADSRWTSPLPNNLNQNSAPLTDVSWIYTLSGTGAKYTLTFDYPGQTGLYFPYKL	353
QY	290	VKAADANNVGLOVKKLANGNVQVEF-----ATSTSANNTANPTAV	331
Dd	354	KTYSD--KVGLQTKLKQAOPVAIQEFSEAAATASAPEATDGROESAEATANEKVNPMFSV	411
QY	332	DETKVAKIVLSGLRFQONTIELSVPTGEGNMKNKVPAMIGNIYLSNNENNAKD	383
Dd	412	NTINAVKVTLSNLKFGCSNTIEFSVPMDDQDNMKNVAPMIGNIYTSSNDANK	463
RESULT	7		
ID	Q49498	PRELIMINARY;	PRT; 702 AA.
OC	AQ498;		
DT	01-NOV-1996	(TrEMBLrel. 01, Created)	
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)	
DE	PMGAL.3	PROTEIN PRECURSOR.	
DE	PMGAL.3		
OS	Mycoplasma gallisepticum.		
CC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;		
OC	Mycoplasmataceae; Mycoplasma.		
OX	NCBI_TaxID=2096;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S6;		
RC	MEDLINE=95010739; PubMed=7925999;		
RA	Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,		
RA	Browning G.F., Whithear K.G., Walker I.D.;		
RT	"the organisation of the multigene family which encodes the major cel-		
RT	surface protein, PMGA, of Mycoplasma gallisepticum.";		
RL	FBS Lett. 352:347-352(1994).		
DR	EMBL; L28424; AAA62417.1; -.		
KW	Signal.		
FT	SIGNAL		
SO	SEQUENCE 702 AA; 75537 MW; 273E8915FE57B9F CRC64;		

Qy	295	ANNVGLQYKLNNGNQVVEF-----ATST-----SAN	322
		:::     :::	:::
Db	365	SSSVALQYSLNKTSSKLINFKPAETVSTNTDQSENEVAFTSTEARSSYKVLVADEAATS	424
Qy	322	NTTANTPTPAVDEIKVAKIVLSGLRFQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNA	381
Db	425	NNEMHTPTVSDINIAKVILSGLTGFGENTIEFSVPEG-----KVAPMIGNMYLTSNSESQ	479
		:::	:::
Qy	382	DKI 384	
Db	480	VKI 482	
RESULT	8		
Q9ZID1		PRELIMINARY; PRT; 645 AA.	
ID	Q9ZID1		
AC	Q9ZID1		
DT	01-MAY-1999	(TEMBLrel. 10, Created)	
DT	01-MAY-1999	(TEMBLrel. 10, Last sequence update)	
DT	01-MAY-1999	(TEMBLrel. 10, Last annotation update)	
DE	M9	PROTEIN.	
DE	M9		
OS	Mycoplasma gallisepticum.		
GN	Mycoplasma gallisepticum.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;		
OC	Mycoplasmataceae; Mycoplasma.		
OX	NCBI_TaxID=2096;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PG31, ATCC19610;		
RX	MEDLINE=99003182; PubMed=9784576;		
RA	Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;		
RT	"A protein (M9) associated with monoclonal antibody-mediated		
RT	agglutination of Mycoplasma gallisepticum is a member of the pmGA		
RT	family.;"		
RL	Infect. Immun. 66:5570-5575(1998).		
DR	EMBL; AF032890; AAC69269.1; -.		
SO	SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAB055 CRC64;		

## RESULT 9

```
Q9ZHR9          PRELIMINARY;      PRT;    644 AA.
AC Q9ZHR9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FBDE5C CRC64;

Query Match      37.2%; Score 745; DB 2; Length 644;
Best Local Similarity 42.3%; Pred. No. 7e-29;
Matches 172; Conservative 65; Mismatches 122; Indels 48; Gaps 11;

Qy 9 DANPNQG-----TQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVNN 61
Db 44 DTNPGDQGMNNAASQELAAARMGLTTVFDSKAKNLGLYVDYKKTQDTLTTRAYDAKTVL 103
Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 104 DNSSSTTNLEAKRLETAIRTAATSKQTFDEQHAELVKVIEELKTTLSNETATLAPYA 163
Qy 122 STAYNQIRNVLVDLYNKASSLITKTLPLNGGTLSDSNEITTANKNINNTL--STINBQK 179
Db 164 AAQYAGIKMHLGSLYDAGKAITTKTLEPEVGDPLTASAVMANTKIVEAIKDEVLPKQK 222
Qy 180 TNADALSNSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV-----NKKY 231
Db 223 ENATKLADSFVKQVLVEKITGVVEAHNS-QPANYSFVGSVDITGTANGQTSIPNNY 281
Qy 232 ARRTVW-NGDEPSSRLANT-----NSTDYSWIYSLAGTNTKYQFSNYPSTGY 282
Db 282 AORTIFTNGDEP--RSVSNTFPVGGTMAQPLSNVSWIYSLAGTAKYTFLEFTYGPSTGY 339
Qy 283 LYFPKLVKAADANNVGLQYKLNNGVQVFEATSTSA-----NNTTANPTPAVDEIKVA 337
Db 340 LYFPKLVNTSDQVKLGLEYKLN-----ATKPSAITFGEQTMNGKTPVNDINVA 391
Qy 338 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 384
Db 392 KVTLANFGSKNIEFSVP-----VEKYSPIGNMNYLSSSPNNWNI 433

RESULT 10
Q49495          PRELIMINARY;      PRT;    650 AA.
AC Q49495; O08060;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HAEMAGGLUTININ PRECURSOR.
GN PMGAL.1.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90714; AAB50152.1; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 650 HAEMAGGLUTININ
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBBB CRC64;

Query Match      37.1%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 7.9e-29;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;

Qy 11 NPNNQG-----TQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVNNN 63
Db 51 NPGDQGMNNAASQELAAARMGLTTVFDSKAKNLGLYVDYKKTQNTLTTRAYDAKTVLDN 110
Qy 64 LNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSST 123
Db 111 SSTTNQNLNEAKRLETAIRTAATSKQTFDEQHAELVKVIEELKTTLSNETATLAPYA 170
Qy 124 AYNQIRNVLVDLYNKASSLITKTLPLNGGTLSDSNEITTANKNINNTL--STINBQKN 181
Db 171 QYAGIKMHLGSLYDAGKAITTKTLEPEVGDPLTASAVMANTKIVEAIKDEVLPKKN 229
Qy 182 ADALSNSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV-----NKKYAR 233
Db 230 ATKLADSFVKQVLVEKITGVVEAHN-RAQANYSFVGSVDITGTANGQTSIPNNY 288
Qy 234 RTVW-NGDEPSSRLANT-----NSTDYSWIYSLAGTNTKYQFSNYPSTGYLY 284
Db 289 RTIFTNGDEP--RSVSNTFPVGGTMAQPLSNVSWIYSLAGTAKYTFLEFTYGPSTGYLY 346
Qy 285 FPKVLVKAADANNVGLQYKLNNGVQVFEATSTANNTTAN-PTPAVDEIKVAKTVLSG 343
Db 347 FPKVLVNTSDQVKLGLEYKLNDA---TEPSAITFGEQTMNGKTPVNDINVAKTVLAN 402
Qy 344 LRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 384
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMNYLSSSPNNWNI 438

RESULT 11
Q49497          PRELIMINARY;      PRT;    649 AA.
AC Q49497;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PMGAL.2 PROTEIN PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RN [1]
RP SEQUENCE FROM N.A.
```

```
RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62416.1; -.
KW SIGNAL.
FT SIGNAL 1 25 POTENTIAL.
SQ SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;

Query Match 37.1%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 8.8e-29;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

Qy 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEVTN 61
Db 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYDYKKTQNTLTAKYDAKTVL 107
Qy 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 108 DNSSSTTONLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
Qy 122 STAYNOIRNNLDVLYNKASSLITKTLDPNGTGLDLSNEITTANKNINNTL--STINEOK 179
Db 168 DAQYAGIRKHLISGLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIVEAIKDEVLNPK 226
Qy 180 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVP-----NKY 231
Db 227 ENATKLADUSFVKQVLKKEITGVEEAHN-KAQPNISFVGYSVDITGTTGTSIPNWDY 285
Qy 232 ARRTYV-NGDEPSSRLANT-----NSTIDVSWIYSLAGTNTKYQFSFSGY 282
Db 286 AQRTIFTNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTGAKYTLEFTYYPGSGY 343
Qy 283 LYFPYKLYKAADANNVGLQYKLNNGVQOQVEFATSTSA-----NNTTANPTPAVDEIKVA 337
Db 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDOTMNGKTPVNDINVA 395
Qy 338 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 384
Db 396 KYTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 437

RESULT 12
ID Q49468 PRELIMINARY; PRT; 647 AA.
AC Q49468; Q53303;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HEMAGGLUTININ HOMOLOG PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93162830; PubMed-8432610;
RA Markham P.F., Glew M.D., Whithear K.G., Walker I.D.;
RT "Molecular cloning of a member of the gene family that encodes pmGA, a
RT hemagglutinin of Mycoplasma gallisepticum.";
RL Infect. Immun. 61:903-909(1993).
DR EMBL; M83178; AAA02996.1; -.
DR EMBL; S55216; AAB25397.2; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 647 HEMAGGLUTININ HOMOLOG..
SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;

Query Match 35.6%; Score 713; DB 2; Length 647;
Best Local Similarity 41.1%; Pred. No. 2.5e-27;
Matches 169; Conservative 60; Mismatches 124; Indels 59; Gaps 12;

Qy 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEVTN 61
Db 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYDYKKTQNTLTAKYDAKTVL 107
Qy 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 108 DNSSSTTONLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
Qy 122 STAYNOIRNNLDVLYNKASSLITKTLDPNGTGLDLSNEITTANKNINNTL--STINEOK 179
Db 168 DAQYAGIRKHLISGLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIVEAIKDEVLNPK 226
Qy 180 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVP-----NKY 231
Db 227 ENATKLADUSFVKQVLKKEITGVEEAHN-KAQPNISFVGYSVDITGTTGTSIPNWDY 285
Qy 232 ARRTYV-NGDEPSSRLANT-----NSTIDVSWIYSLAGTNTKYQFSFSGY 282
Db 286 AQRTIFTNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTGAKYTLEFTYYPGSGY 343
Qy 283 LYFPYKLYKAADANNVGLQYKLNNGVQOQVEFATSTSA-----NNTTANPTPAVDEIKVA 337
Db 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDOTMNGKTPVNDINVA 395
Qy 338 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 384
Db 396 KYTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 437

RESULT 13
ID Q9KH15 PRELIMINARY; PRT; 656 AA.
AC Q9KH15;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ADHESIN PMGAL.2.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmGA multigene family of Mycoplasma
RT gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91413.1; -.
SQ SEQUENCE 656 AA; 70875 MW; 58EA7E075FC617E1 CRC64;

Query Match 35.1%; Score 704; DB 2; Length 656;
Best Local Similarity 41.0%; Pred. No. 6.9e-27;
Matches 166; Conservative 67; Mismatches 124; Indels 48; Gaps 12;

Qy 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEVTN 61
Db 63 DTPGCGGMMNAASQELAAARMGLTTTFDSKAKNLGLYDYKKTQNTLTAKYDAKTVL 122
Qy 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 123 QNSSATLEQVKNTSALQTAINTANSNKQKQDQDHSNLLMSYKMLMATLAKKETAVMTLK 182
Qy 122 STAYNOIRNNLDVLYNKASSLITKTLDPNGTGLDLSNEITTANKNINNTL--TINEOK 179
Db 183 DPKYSAILDQINGVSKGEEVQHTLDVPS-GIVPAANTITEITKIEVISEKTLQDQK 241
Qy 180 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVP----- 227
Db 242 NNADQFANYQSTFKDKLENVEDA-----KKMGOPANYSFVGYSVDVTGTSQGETTIP 295
Qy 228 NYKYARRTYVNGDEPSSRLANTNS-----ITDVSWIYSLAGTNTKYQFSFSGY 279
```

```
Db 296 WNFQAIAITSGNQPKVATTTGGEDQSTAKPLSDVSWIYSLAGTGAKTLEFTYGPS 355
QY 280 TGYLYFFPKLVKAADANNVGLQYKLN--NGNVQVVEFATSTSANNTTPPAVDEIKVAK 338
Db 356 TGYLYFFPKLVKAND--DVGLOYLKLSNEILTPPIFGECT-----TTNGPAAVTENINAK 409
QY 339 IVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNENNADK 383
Db 410 VRLTGLAFGKNTIEFSVP-----MSKVAPMIGNMYITSSDTETNK 449

RESULT 14
ID 005122 PRELIMINARY; PRT; 703 AA.
AC 005122;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HAEMAGGLUTININ.
GN PMGAL.9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Giew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RT Browning G.F., Whithar K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U90714; AAB50154.1;
DR InterPro: IPR002819; HD.
SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73CBC5 CRC64;

Query Match 34.5%; Score 690.5; DB 2; Length 703;
Best Local Similarity 35.0%; Pred. No. 3.4e-26;
Matches 167; Conservative 75; Mismatches 120; Indels 115; Gaps 12;

QY 2 CMSITTKDANP-----NNGQT-----QLEAARMELTDLINAK 33
Db 26 CTSATIPTLNPENPEKPDMPNPPSGMGGNTNPGMDTAQELASAKAALTTLNRE 85
QY 34 AMTLASLDYAKIEASLSAYSEATVNNLNATLEOLKMAKTNLESAINQANTDKTFD 93
Db 86 SEKVLGVVDYAKIKADITSAYTAKTSDSTSLVQVKTATSKLQTAIDKAASDKOFE 145
QY 94 NEHPNLVEAYKALKTTLEQR-ATNLEGLSTAYNOIRNVLNLYNKASSLITKLDPLNG 152
Db 146 QDHKDLLMPSEKLTLSQKNATVL--LNQPKYSAILINKINSIYAQGEVIVIRLDPVS- 202
QY 153 GTLLDSNEITTANKNNINTLS--TINQKTNADALSNSFIKKVQNNEQSFVGTFTNANV 210
Db 203 GAIPTAASITKVNDINKAISENLKPKDNADAFANYQFFKL--DKTKIMGSTNMWK 259
QY 211 QPSNYSFVAESADVTPV-----NKKYARTVWNGDEPSSRLANTNS----- 252
Db 260 QPQNYSEFVGSVGTGMSQGTIPNNWFAQRIWSSGAPRAPLASOTETPQAETPPWSA 319
QY 253 -----ITDYSWIYSLA 263
Db 320 PQGVEPAQQQDSSPKQASQEVSPPAEVOQAQADTEQPATSQGTPLTDVSWIYSL 379
QY 264 GTNTKYQFSFNYPSTGYLYFFPKLVKAADANNVGLQYKLNNGNVQVVEFATSTSANNT 323
```

```
Db 380 GTDVKYFTTFNYFGPSMAYLYFPYKLVKSD--SVGLQYKLNPNVAFNFGSETNAN-- 435
QY 324 TANTPPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNENN 380
Db 436 --GPAASVDNINAKNLANLNFGEITIEFSVP-----MNKVAPMIGNMYITSDVAN 485

RESULT 15
ID 09L8D5 PRELIMINARY; PRT; 419 AA.
AC 09L8D5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PMGA-LIKE PROTEIN 9.3 (FRAGMENT).
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.T., Branton S.B., Hanson L.A., Minion F.C., Lott B.D.,
RT May J.D., Hughlett M.B.;
RT "A novel pmga-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF210770; AAF29525.1;
FT NON_TER 419
SQ SEQUENCE 419 AA; 45072 MW; FE5EBE37F2DB3B0C CRC64;

Query Match 30.98; Score 619.5; DB 2; Length 419;
Best Local Similarity 40.1%; Pred. No. 5.4e-23;
Matches 149; Conservative 58; Mismatches 122; Indels 43; Gaps 11;

QY 9 DANPNNGQT-----QLEAARMELTDLINAKMTLASLDYAKIEASLSAYSEATVN 61
Db 61 DINPGGQGNMDSAAQELTAARTALTSLLASKNANIEMYSYAKIQNTLIAAYTTAEQTS 120
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 121 QNSSATLEQVKNATSALQTAINTANSNKKQKQPDQHSNLLMSYKNLMATLAKKTTVMTLK 180
QY 122 STAYNOIRNVLNLYNKASSLITKLDPLNGTLLDSNEITTANKNNINTLS--TINQK 179
Db 181 DPKYSAILDQINGVSCKGEELVQHTLDPVS--GIVPAANTITEETKIEEIVSEKTLQDOK 239
QY 180 TNADALSN--SFI--KKVIONNEQSFVGTFTNANVQPSNYSFVAESADVTPV----- 227
Db 240 NNADQFDNYQSFTLDKTKLENVEDA-----KKMGQPAKNSYFVGSVDVTGTSQETIP 293
QY 228 NYKYARTVWNGDEPSSRLANTNS-----ITDYSWIYSLAGTNTKYQFSFNYPST 279
Db 294 WNFQAIAITSGNQPKVATTTGGEDQSTAKPLSDVSWIYSLAGTGAKTLEFTYGPS 353
QY 280 TGYLYFFPKLVKAADANNVGLQYKLN--NGNVQVVEFATSTSANNTTPPAVDEIKVAK 338
Db 354 TGYLYFFPKLVKAND--DVGLOYLKLSNEILTPPIFGECT-----TTNGPAAVTENINAK 407
QY 339 IVLSGLRFGONT 350
Db 408 VRLTGLAFGKNT 419

Search completed: June 12, 2002, 10:50:42
Job time: 205 sec
```

